



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 166995

TO: Manjunath N Rao
Location: rem-2a01/2c70
Art Unit: 1652
Friday, September 23, 2005
Case Serial Number: 10/777828

From: Toby Port
Location: Biotech-Chem Library
REM-1A59
Phone: 571-272-2523
toby.port@uspto.gov

Search Notes

Examiner Rao,

See attached results.

If you have any questions about this search feel free to contact me at any time.

Thank you for using STIC search services!

Toby Port
X22523

This Page Blank (uspto)

STIC-Biotech/ChemLib

165996

From: Rao, Manjunath N.
Sent: Friday, September 16, 2005 2:37 PM
To: STIC-Biotech/ChemLib
Subject: 10777828

From: Manjunath N. Rao
Art Unit 1652, Room 2A01
Mail Box in Room 2C70
Phone: 272-0939

Date: 9-16-05

Please search the following as soon as possible for application with serial number

10/777828

1. SEQ ID NO: 8 AND Nucleotides 1-115 and nucleotides 428-1011 of SEQ ID NO:8 against all commercial nucleic acid sequence databases, issued patents/published applications nucleic acid sequence database and pending application nucleic acid sequence database. Please provide a print of all results

If you have any questions please call me at the above phone number.

Thanks

9/16/2005

Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2005, 07:38:28 ; Search time 4623.4 Seconds
(without alignments)
10595.706 Million cell updates/sec

Title: US-10-777-828-8
Perfect score: 1011
Sequence: 1 ccacctcgcctcttagcat.....aagattgctgcgtgtctga 1011

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_to.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match	Description
1	1011	100.0	BD228354
2	1006.2	99.5	E38420 Novel poly
3	1006.2	99.5	AF064860 Homo sapi
4	1006.2	99.5	AL163280 Homo sapi
5	985.4	97.5	AL954227 Pan trogl
6	978.4	96.8	AB041416 Homo sapi
7	957	94.7	AB041415 Pan panis
8	955.6	94.5	AB041413 Homo sapi
9	948.4	93.8	AB041412 Homo sapi
10	945.6	93.5	AB041411 Gorilla g
11	930.2	92.0	CQ731786 Sequence
12	928.8	91.9	AJ006078 Homo sapi
13	928.6	91.8	AB020337 Homo sapi
14	928.6	91.8	E38419 Novel poly
15	928.6	91.8	AY372061 Homo sapi
16	928.2	91.8	AF145784 Homo sapi
17	926.4	91.6	AB041417 Pongo pyg
18	771.8	76.3	AY231145 Macaca mu
19	586	58.0	AC150794 Bos tauru

20	532.4	52.7	196900	2	AC020851	AC020851 Mus muscu	
21	527.6	52.2	149964	2	AC120145	AC120145 Mus muscu	
c	527.6	52.2	186956	2	AC120346	AC120346 Mus muscu	
23	525.6	52.0	4933	10	BC057887	BC057887 Mus muscu	
24	525.6	52.0	5069	10	BC051669	BC051669 Mus muscu	
25	522.4	51.7	927	10	AF254738	AF254738 Mus muscu	
c	323.8	32.0	65400	2	AC109264	AC109264 Mus muscu	
27	269	26.6	97702	2	AC151319	AC151319 Xenopus t	
c	209.4	20.7	226720	2	BCX93631	BCX93631 Danio rer	
29	206.2	20.4	2317	5	BC066477	BC066477 Danio rer	
30	185.2	18.3	917	10	AB039137	AB039137 Mus muscu	
31	185.2	18.3	917	10	AB039140	AB039140 Mus muscu	
32	185.2	18.3	917	10	AB039141	AB039141 Mus muscu	
33	185.2	18.3	917	10	AB039142	AB039142 Mus muscu	
34	183.6	18.2	917	10	AB039136	AB039136 Mus muscu	
35	183.6	18.2	917	10	AB039138	AB039138 Mus muscu	
36	183.6	18.2	917	10	AB039143	AB039143 Mus spici	
37	182	18.0	917	10	AB039134	AB039134 Mus muscu	
38	182	18.0	917	10	AB039135	AB039135 Mus muscu	
39	182	18.0	917	10	AB039139	AB039139 Mus muscu	
40	182	18.0	1175	10	AF029790	AF029790 Mus muscu	
41	182	18.0	188273	10	BX936296	BX936296 Mouse DNA	
c	42	182	18.0	219368	10	AC098721	AC098721 Mus muscu
c	43	181.2	17.9	207679	2	BX511272	BX511272 Danio rer
c	44	181	17.9	110000	2	AC096079	Continuation (2 of
c	45	181	17.9	231235	2	AC136678	AC136678 Rattus no

ALIGNMENTS

RESULT 1
BD228354
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
ORIGIN

BD228354
UDP-galactose: beta-N-acetyl-glucosamine beta 1,3
galactosyltransferases, beta 3 Gal-T5.
BD228354.1 GI:33038124
JP 2002530071-A/8.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Clausen.H. and Amado.M.
UDP-galactose: beta-N-acetyl-glucosamine beta 1,3
galactosyltransferases, beta 3 Gal-T5
Patent: JP 2002530071-A 8 17-SEP-2002;
HENRIK CLAUSEN
OS Homo sapiens (human)
PN JP 2002530071-A/8
PD 17-SEP-2002
PF 11-NOV-1999 JP 2000582542
PR 13-NOV-1998 DK PA 199801483
PI HENRIK CLAUSEN, MARGARIDA AMADO
PC C12N15/09, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12N9/10, C12N15/10, C12N5/00
CC UDP-galactose: beta-N-acetyl-glucosamine beta 1,3 CC
galactosyltransferases,
beta 3 Gal-T5
FH Key Location/Qualifiers
FT CDS Location/Qualifiers
1. 1011
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

Query Match 100.0%; Score 1011; DB 6; Length 1011;
Best Local Similarity 100.0%; Pred. No. 7.8e-270;
Matches 1011; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCACCTCAGCCCTCTAGCATAAACTAGACATCCTCATGCTTTTGGGCTCTAATCATT 60
Db 1 CCACCTCAGCCCTCTAGCATAAACTAGACATCCTCATGCTTTTGGGCTCTAATCATT 60
QY 61 GGATTTTGTCTCTTTCAGATGCTTTCCGGAAGATGAGATTGATATATCTGCCCTCTG 120
Db 61 GGATTTTGTCTCTTTCAGATGCTTTCCGGAAGATGAGATTGATATATCTGCCCTCTG 120
QY 121 GTTCTGGGGCTCTTTGTTGTTATTTTAGCATGTACATCTAAATCTCTTCAAAGAACAG 180
Db 121 GTTCTGGGGCTCTTTGTTGTTATTTTAGCATGTACATCTAAATCTCTTCAAAGAACAG 180
QY 181 TCCTTTGTTTACAAGAAAGAGCGGAATCTCTTAAGCTCCAGATACAGACTGCAGGCAG 240
Db 181 TCCTTTGTTTACAAGAAAGAGCGGAATCTCTTAAGCTCCAGATACAGACTGCAGGCAG 240
QY 241 ACACCTCCCTCTCTGCTGCTGCTGCTCCTCCACAAACAGTTGGCTGAGCGCATG 300
Db 241 ACACCTCCCTCTCTGCTGCTGCTGCTCCTCCACAAACAGTTGGCTGAGCGCATG 300
QY 301 GCATCCGGCAGACGTGGGGGAAAGAGAGGACGGTGAAGGGAAGCAGCTGAAGACATTC 360
Db 301 GCATCCGGCAGACGTGGGGGAAAGAGAGGACGGTGAAGGGAAGCAGCTGAAGACATTC 360
QY 361 TTCTCTCTGGGACCAACAGCAGTGCAGCGGAACAAAGAGGTTGGACCGAGAGCCAG 420
Db 361 TTCTCTCTGGGACCAACAGCAGTGCAGCGGAACAAAGAGGTTGGACCGAGAGCCAG 420
QY 421 CGACACGGGGACATTTATCCAGAAGGATTTCTTAGACGTCTATTACAATCTGACCCCTGAAG 480
Db 421 CGACACGGGGACATTTATCCAGAAGGATTTCTTAGACGTCTATTACAATCTGACCCCTGAAG 480
QY 481 ACCATGATGGGCAATGATGAGTGGTCCATCGCTTTTGTCTCAGCGGGCGTTTGTGATGAAA 540
Db 481 ACCATGATGGGCAATGATGAGTGGTCCATCGCTTTTGTCTCAGCGGGCGTTTGTGATGAAA 540
QY 541 ACAGACTCAGACATGTTTCAATCAATGTTGACTATCTGACTGAACTCTCTGAAGAAAC 600
Db 541 ACAGACTCAGACATGTTTCAATCAATGTTGACTATCTGACTGAACTCTCTGAAGAAAC 600
QY 601 AGAACACACAGGTTTTTCACTGGCTTTCTTGAACACTCAATGATTTCCCATCAGGCAGCCA 660
Db 601 AGAACACACAGGTTTTTCACTGGCTTTCTTGAACACTCAATGATTTCCCATCAGGCAGCCA 660
QY 661 TTCAGCAAGTGGTTTGTTCAGTAAATCTGAATATCCGTGGGACAGGTACCCACATTTGCG 720
Db 661 TTCAGCAAGTGGTTTGTTCAGTAAATCTGAATATCCGTGGGACAGGTACCCACATTTGCG 720
QY 721 TCCGGCACGGCTACGTGTTTCTGGCGACGTGGCGAGTCAAGTGTACAATGTCTCCAAAG 780
Db 721 TCCGGCACGGCTACGTGTTTCTGGCGACGTGGCGAGTCAAGTGTACAATGTCTCCAAAG 780
QY 781 AGCGTCCCATACATTTAACTGGAAGACGTGTTTGTGGGGCTCTGCTCGAAGGCTGAAC 840
Db 781 AGCGTCCCATACATTTAACTGGAAGACGTGTTTGTGGGGCTCTGCTCGAAGGCTGAAC 840
QY 841 ATCAGATGGAGGAGCTCCATCCAGCGACCTTTTTCAGGGGGCTTACGCTTTCTCC 900
Db 841 ATCAGATGGAGGAGCTCCATCCAGCGACCTTTTTCAGGGGGCTTACGCTTTCTCC 900
QY 901 GTATGCTCTTTCAGAGGATCGTGGCCCTGCCACTTTCATCAAGCCCTCGGACTCTCTTGAC 960
Db 901 GTATGCTCTTTCAGAGGATCGTGGCCCTGCCACTTTCATCAAGCCCTCGGACTCTCTTGAC 960
QY 961 TACTGGCAGGCTCTAGAGAAATTCGCGGGGGAAGATTGTCCGCTCTGTCTGA 1011
Db 961 TACTGGCAGGCTCTAGAGAAATTCGCGGGGGAAGATTGTCCGCTCTGTCTGA 1011

RESULT 2
E38420
LOCUS
DEFINITION Novel polypeptide.
10562 bp DNA linear PAT 31-JAN-2002

E38420
E38420.1 GI:18626994
JP 2000245464-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 10562)
AUTHORS Narimatsu, H., Isshiki, S., Togayauchi, A. and Sasaki, K.
TITLE Novel polypeptide
JOURNAL Patent: JP 2000245464-A 2 12-SEP-2000;
KYOWA HAKKO KOGYO CO LTD
COMMENT OS Homo sapiens (human)
PN JP 2000245464-A/2
PD 12-SEP-2000
PF 25-FEB-1999 JP 1999047571
PR HISASHI NARIMATSU, SOICHIRO ISSHIKI, AKIRA TOGAYAUCHI, PI
KATSUTOSHI SASAKI
PC C12N15/09, A01K67/027, C12N1/21, C12N5/10, C12N9/10, C12P19/00, PC
C12P21/02,
PC C12P21/08, C12Q1/68, G01N33/53// (C12N1/21, C12R1:185), (C12N5/10,
C12R1:91),
PC (C12P21/02, C12R1:185), (C12P21/02, C12R1:91), C12N15/00, C12N5/00,
PC (C12N5/00, C12R1:91)
CC
FH Key Location/Qualifiers
FT promoter (1)..(5000)
FT exon (5001)..(5140)
FT exon (5001)..(5273)
FT exon (5459)..(5567)
FT exon (7427)..(7586)
FT exon (8234)..(10562).
FEATURES
source
1..10562
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 99.5%; Score 1006.2; DB 6; Length 10562;
Best Local Similarity 99.7%; Pred. No. 1.9e-268;
Matches 1008; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CCACCTCAGCCCTCTAGCATAAACTAGACATCCTCATGCTTTTGGGCTCTAATCATT 60
Db 8156 CCACCTCAGCCCTCTAGCATAAACTAGACATCCTCATGCTTTTGGGCTCTAATCATT 8215
QY 61 GGATTTTGTCTCTTTCAGATGCTTTCCGGAAGATGAGATTGATATATCTGCCCTCTG 120
Db 8216 GGATTTTGTCTCTTTCAGATGCTTTCCGGAAGATGAGATTGATATATCTGCCCTCTG 8275
QY 121 GTTCTGGGGCTCTTTGTTGTTATTTTAGCATGTACAGTCTAAATCTCTTCAAAGAACAG 180
Db 8276 GTTCTGGGGCTCTTTGTTGTTATTTTAGCATGTACAGTCTAAATCTCTTCAAAGAACAG 8335
QY 181 TCCTTTGTTTACAAGAAAGAGCGGAACTTCCTTAAGCTCCAGATACAGACTGCAGGCAG 240
Db 8336 TCCTTTGTTTACAAGAAAGAGCGGAACTTCCTTAAGCTCCAGATACAGACTGCAGGCAG 8395
QY 241 ACACCTCCCTCTCTGCTGCTGCTGCTCCTCCACAAACAGTTGGCTGAGCGCATG 300
Db 8396 ACACCTCCCTCTCTGCTGCTGCTGCTCCTCCACAAACAGTTGGCTGAGCGCATG 8455
QY 301 GCATCCGGCAGACGTGGGGGAAAGAGAGGACGGTGAAGGGAAGCAGCTGAAGACATTC 360
Db 8456 GCATCCGGCAGACGTGGGGGAAAGAGAGGACGGTGAAGGGAAGCAGCTGAAGACATTC 8515
QY 361 TTCTCTCTGGGACCAACAGCAGTGCAGCGGAACAAAGAGGTTGGACCGAGAGCCAG 420
Db 8516 TTCTCTCTGGGACCAACAGCAGTGCAGCGGAACAAAGAGGTTGGACCGAGAGCCAG 8575
QY 421 CGACACGGGGACATTTACCAAGAGATTTCCTAGACGTCTATTACAAATCTGACCCCTGAAG 480

```
Db 8576 CGACGCGGACATTTATCCAGAAAGATTTCCTAGACGCTATTATCAATCTGACCCCTGAAG 8635
Qy 481 ACCATGATGGGCATAGATGGGTCCATCGCTTTTCTCTCAGGCGCGCTTTGTGATGAAA 540
Db 8636 ACCATGATGGGCATAGATGGGTCCATCGCTTTTCTCTCAGGCGCGCTTTGTGATGAAA 8695
Qy 541 ACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGAACTGCTCTTCTGAAGAAAAAC 600
Db 8696 ACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGAACTGCTCTTCTGAAGAAAAAC 8755
Qy 601 AGAACACACAGTCTTCTCAGTGGCTCTTGAACCTCAATGATGTTTCCCATCAGGCGGACCA 660
Db 8756 AGAACACACAGTCTTCTCAGTGGCTCTTGAACCTCAATGATGTTTCCCATCAGGCGGACCA 8815
Qy 661 TTCAGCAAGTGTGTTGTCAGTAAATCTCAATATCCGTGGGACAGTACCCACCATTTCTGC 720
Db 8816 TTCAGCAAGTGTGTTGTCAGTAAATCTCAATATCCGTGGGACAGTACCCACCATTTCTGC 8875
Qy 721 TCCGCGACCGGCTACGTTGTTTCTGGCGACGTGGCGAGTCAAGTGTACAATGTTCTTCCAAAG 780
Db 8876 TCCGCGACCGGCTACGTTGTTTCTGGCGACGTGGCGAGTCAAGTGTACAATGTTCTTCCAAAG 8935
Qy 781 AGCGTCCCATCATTAATACTGAAGACGTGTTTGTGGGGCTCTGCTCGAAGGCTGAAC 840
Db 8936 AGCGTCCCATCATTAATACTGAAGACGTGTTTGTGGGGCTCTGCTCGAAGGCTGAAC 8995
Qy 841 ATCAGATTGGAGGAGCTCCACTCCAGCGGACCTTTTTCAGGCGGCTTAGGCTTCTCC 900
Db 8996 ATCAGATTGGAGGAGCTCCACTCCAGCGGACCTTTTTCAGGCGGCTTAGGCTTCTCC 9055
Qy 901 GTATGCTCTTTTCAGGAGGATCGTGCCCTGCCACTTCATCAAGCCCTCGGAGCTCTCTTGAC 960
Db 9056 GTATGCTCTTTTCAGGAGGATCGTGCCCTGCCACTTCATCAAGCCCTCGGAGCTCTCTTGAC 9115
Qy 961 TACTGGCAGGCTCTAGAGAAATCCCGGGGGGAGAGATTGTGCGCCTGTCTGA 1011
Db 9116 TACTGGCAGGCTCTAGAGAAATCCCGGGGGGAGAGATTGTGCGCCTGTCTGA 9166

RESULT 3
AF064860
LOCUS Homo sapiens chromosome 21 clone PAC 70124 map 21q22.3, complete
DEFINITION
ACCESSION AF064860
VERSION AF064860.2 GI:18958624
KEYWORDS HTG; HTGS DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T.,
Park,H.-S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.-K., Soeda,E.,
Ohki,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K.,
Polley,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R.,
Patterson,D., Reichwald,K., Rump,A., Schillhabel,M.B., Schudy,A.,
Zimmermann,W., Rosenthal,A., Kudoh,J., Kawasaki,K., Asakawa,S.,
Shantani,A., Sasaki,T., Nagamine,K., Mitsuyma,S.,
Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordsiek,G.,
Hornischer,K., Brand,P., Scharfe,M., Schoen,O., Desario,A.,
Reichelt,J., Kauer,G., Bloeker,H., Ramser,J., Beck,A., Klages,S.,
Hennig,S., Riessmann,L., Dagand,E., Haaf,T., Wehrmeyer,S.,
Horzym,K., Gardiner,K., Nizetic,D., Francis,F., Lehrach,H.,
Reinhardt,R. and Yaspo,M.Laure.
The DNA sequence of human chromosome 21
Nature 405 (6784), 311-319 (2000)
20289799
10830953
REFERENCE
JOURNAL
MEDLINE
PUBMED
2 (bases 1 to 170121)
AUTHORS Rump,A., Dagand,E., Hildmann,T., Nordsiek,G., Drescher,B.,
Weber,J., Schattevoy,R., Yaspo,M.-L. and Rosenthal,A.
```

```
TITLE Direct Submission
JOURNAL Submitted (12-MAY-1998) Genome Analysis, Institute of Molecular
REFERENCE Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
AUTHORS 3 (bases 1 to 170121)
Rump,A., Dagand,E., Hildmann,T., Nordsiek,G., Drescher,B.,
Weber,J., Schattevoy,R., Yaspo,M.-L., Rosenthal,A., Yaspo,M.-L. and
Rosenthal,A.
TITLE Direct Submission
JOURNAL Submitted (27-FEB-2002) Genome Analysis, Institute of Molecular
REFERENCE Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
AUTHORS 4 (bases 1 to 170121)
Rump,A., Dagand,E., Hildmann,T., Nordsiek,G., Drescher,B.,
Weber,J., Schattevoy,R., Yaspo,M.-L. and Rosenthal,A.
TITLE Direct Submission
JOURNAL Submitted (05-MAR-2002) Genome Analysis, Institute of Molecular
REFERENCE Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
COMMENT On Feb 27, 2002 this sequence version replaced gi:3171153.
FEATURES
source
1. 170121
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
/clone="PAC 70124"
ORIGIN
Query Match 99.5%; Score 1006.2; DB 9; Length 170121;
Best Local Similarity 99.7%; Pred. No. 2.1e-268;
Matches 1008; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 CCACCTCAGCCTCCTAGCATAAACTAGACACATCCTCATGCTTTTGGAGTCTAATCATTT 60
Db 89970 CCACCTCAGCCTCCTAGCATAAACTAGACACATCCTCATGCTTTTGGAGTCTAATCATTT 90029
Qy 61 GGATTTTGTCTTTCAGATGGCTTTCCGGAAGATGAGATTGATGATATCTGCTCTG 120
Db 90030 GGATTTTGTCTTTCAGATGGCTTTCCGGAAGATGAGATTGATGATATCTGCTCTG 90089
Qy 121 GTTCTGGGGGCTCTTGTGTTGTTATTTTAGCATGTACAGTCTAAATCCTTTTCAAAGAACAG 180
Db 90090 GTTCTGGGGGCTCTTGTGTTGTTATTTTAGCATGTACAGTCTAAATCCTTTTCAAAGAACAG 90149
Qy 181 TCCCTTTGTTTACAAAGAAAGACGGAACTTCTTAAAGCTCCAGATACAGACTGCGAGGACAG 240
Db 90150 TCCCTTTGTTTACAAAGAAAGACGGAACTTCTTAAAGCTCCAGATACAGACTGCGAGGACAG 90209
Qy 241 ACACCTCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
Db 90210 ACACCTCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 90269
Qy 301 GCCATCCGGCAGCGTGGGGGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Db 90270 GCCATCCGGCAGCGTGGGGGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 90329
Qy 361 TTCTCTCTGGGGGACACACAGCAGCTGCGAGCGGAAACAAAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Db 90330 TTCTCTCTGGGGGACACACAGCAGCTGCGAGCGGAAACAAAGAGAGAGAGAGAGAGAGAGAGAGAG 90389
Qy 421 CGACACGGGGACATATTCAGAAAGATTTCCTAGACGCTTATTAATCAATCTGACCCCTGAAG 480
Db 90390 CGACACGGGGACATATTCAGAAAGATTTCCTAGACGCTTATTAATCAATCTGACCCCTGAAG 90449
Qy 481 ACCATGATGGGCATAGATGGGTCCATCGCTTTTCTCTCAGGCGCGCTTTGTGATGAAA 540
Db 90450 ACCATGATGGGCATAGATGGGTCCATCGCTTTTCTCTCAGGCGCGCTTTGTGATGAAA 90509
Qy 541 ACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGAATCTGCTTCTGAAGAAAAAC 600
Db 90510 ACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGAATCTGCTTCTGAAGAAAAAC 90569
Qy 601 AGAACACACAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
```

```
Db 90570 AGAACCAACGAGGTTTTTCACTGGCTTCTTGAACCTCAATGAGTTTCCCATCAGGCAGCCA 90629
Qy 661 TTCAGCAAGTGGTTTGTTCAGTAAATCTCAATATCGTGGGACAGTACCCACCACTTCTGC 720
Db 90630 TTCAGCAAGTGGTTTGTTCAGTAAATCTCAATATCGTGGGACAGTACCCACCACTTCTGC 90689
Qy 721 TCCGGCACCGGCTACGTTCTTTCTGGCACCTGGCGAGTCAGGTGTACAATGTCTCCAAG 780
Db 90690 TCCGGCACCGGCTACGTTCTTTCTGGCACCTGGCGAGTCAGGTGTACAATGTCTCCAAG 90749
Qy 781 AGCGTCCCATCAATTAACCTGAAGACGTGTTGTGGGGCTCTCCCTCGAAGGCTGAAC 840
Db 90750 AGCGTCCCATCAATTAACCTGAAGACGTGTTGTGGGGCTCTCCCTCGAAGGCTGAAC 90809
Qy 841 ATCAGATTGGAGGAGCTCCATCCCGACCGACCTTTTTCAGGGGGCTTAGGCTTCTCC 900
Db 90810 ATCAGATTGGAGGAGCTCCATCCCGACCGACCTTTTTCAGGGGGCTTAGGCTTCTCC 90869
Qy 901 GTATGCTCTTTCAGGAGGATCGTGGCCCTGCCACTTCATCAAGCCTCGGACTCTCTTGGAC 960
Db 90870 GTATGCTCTTTCAGGAGGATCGTGGCCCTGCCACTTCATCAAGCCTCGGACTCTCTTGGAC 90929
Qy 961 TACTGGCAGGCTCTAGAGAATTCGCGGGGGAAGATTGTCGCGCTGTCTGA 1011
Db 90930 TACTGGCAGGCTCTAGAGAATTCGCGGGGGAAGATTGTCGCGCTGTCTGA 90980

RESULT 4
HS21C080 340000 bp DNA linear PRI 24-MAY-2000
LOCUS Homo sapiens chromosome 21 segment HS21C080.
DEFINITION AL163280 AP001735 BA000005
ACCESSION AL163280.2 GI:7717369
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 340000)
Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T.,
Park,H.S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.K., Soeda,E.,
Ohki,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K.,
Polley,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R.,
Patterson,D., Reichwald,K., Rump,A., Schillhabel,M., Schudy,A.,
Zimmermann,W., Rosenthal,A., Kudon,J., Shibuya,K., Kawasaki,K.,
Asakawa,S., Shintani,A., Sasaki,T., Nagamine,K., Mitsuyama,S.,
Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordtiek,G.,
Hornischer,K., Brandt,P., Scharfe,H., Schoen,O., Desario,A.,
Reichert,J., Kauer,G., Bloecker,H., Ramser,J., Beck,A., Klages,S.,
Hennig,S., Rieselmann,L., Dagand,E., Wehrmeyer,S., Borzym,K.,
Gardiner,K., Nizetic,D., Francis,F., Lehrach,H., Reinhardt,R. and
Yaspo,M.L.
Direct Submission
Submitted (05-MAY-2000) The Chromosome 21 Mapping and Sequencing
Consortium: * RIKEN Genomic Sciences Center, Human Genome Research
Group * Institute of Molecular Biotechnology, Genome Analysis *
Keio University School of Medicine, Dept. of Molecular Biology *
GBF, Dept. of Genome Analysis * Max-Planck Institute for Molecular
Genetics (addresses see below)
The Chromosome 21 Mapping and Sequencing Consortium consists of
* RIKEN Genomic Sciences Center, Human Genome Research Group, *
Sagamihara 228-8555, Japan,
* e-mail: sakaki@gsic.riken.go.jp
* URL: http://hgp.gsic.riken.go.jp/
and
* Institute of Molecular Biotechnology, Genome Analysis, *
Butenbergrasse 11, D-07745 Jena, Germany,
* e-mail: gscj-submit@genome.imb-jena.de
* URL: http://genome.imb-jena.de/
and
* Keio University School of Medicine, Dept. of Molecular Biology, *
Tokyo 160-8582, Japan,
* e-mail: shimizu@dm-med.keio.ac.jp
```

```
FEATURES
source
1. 340000
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
<1..125946
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
/clone="P160L9, 5' partial"
/clone_lib="RPC11.3-5 PAC library"
/note="Accession No. AF121897"
73410..243533
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
/clone="P70124"
/clone_lib="RPC11.3-5 PAC library"
/note="Accession No. AF064860"
224137..>340000
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
/clone="P206A10, 3' partial"
/clone_lib="RPC11.3-5 PAC library"
/note="Accession No. AF121782"
286628..>340000
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
/clone="BAC-291B3, 3' partial"
/clone_lib="BAC library"
/note="Accession No. AF064857"
1088..1566
/note="LIM4"
/rpt_family="LINE/L1"
/rpt_type=DISPERSED
2398..2495
/note="(TCTA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
complement(2496..2878)
/note="THE1C"
/rpt_family="LTR/MaLR"
/rpt_type=DISPERSED
2879..2931
/note="(TCTA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
2932..3004
/note="(TCCA)n"
/rpt_family="Simple_repeat"

repeat_region
repeat_region
repeat_region
repeat_region
repeat_region
```

```
* URL: http://adenine.dmb.med.keio.ac.jp/
and
* GBF, Dept. of Genome Analysis,
Mascheroder Weg 1, D-38124 Braunschweig, Germany, * e-mail:
info.genome@bf.de
* URL: http://genome.gbf.de/
and
* Max-Planck Institute for Molecular Genetics,
Innesstrasse 73, D-14195 Berlin, Germany,
* e-mail: info-chr21@molgen.mpg.de
* URL: http://chr21.rz-berlin.mpg.de/.
```

```
repeat_region      /rpt_type=TANDEM
3005..3074
/note="TCCA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
3473..19157
/gene="SH3BGR"
join(<3473..3565,12649..12678,15331..15460,18739..19157)
/gene="SH3BGR"
CDS
join(3473..3565,12649..12678,15331..15426)
/gene="SH3BGR"
/note="Accession No. X93498"
/codon_start=1
/product="21-Glutamic Acid Rich protein 21-GARP"
/protein_id="CAB90445.1"
/db_xref="GI:7717370"
/db_xref="GOA:P55822"
/db_xref="UniProt/Swiss-Prot:P55822"
translation="GSEKAEKGGETEAQKGESEVDGNLPRAQEKNEBEGETATETEE
IAMEGAEEAESEETAEGERPGEDEDS"
3473..3565
/gene="SH3BGR"
/number=4
5542..5829
/note="AluSx"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(6887..7067)
/note="L1MD3"
/rpt_family="LINE/L1"
/rpt_type=DISPERSED
complement(8482..8762)
/note="AluJo"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(8865..8964)
/note="L2"
/rpt_family="LINE/L2"
/rpt_type=DISPERSED
9131..9424
/note="AluSq"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(9723..9882)
/note="L2"
/rpt_family="LINE/L2"
/rpt_type=DISPERSED
complement(9989..10077)
/note="L2"
/rpt_family="LINE/L2"
/rpt_type=DISPERSED
complement(10078..10429)
/note="THE1B"
/rpt_family="LTR/MaLR"
/rpt_type=DISPERSED
complement(10430..10651)
/note="L2"
/rpt_family="LINE/L2"
/rpt_type=DISPERSED
10980..11000
/note="(TTTA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
complement(11002..11282)
/note="AluSp"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(11315..11556)
/note="AluJo"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
11702..11753
/note="MIR"

/rpt_family="SINE/MIR"
/rpt_type=DISPERSED
12649..12678
/gene="SH3BGR"
/number=5
complement(13115..13425)
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
complement(13598..13892)

Query Match      99.5%; Score 1006.2; DB 9; Length 340000;
Best Local Similarity 99.7%; Pred. No. 2.2e-269;
Matches 1008; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCACCTCAGCCTCTAGCATATAAACTAGACACATCCTCATCTTTTGAGGTCTAATCATTT 60
DB 163378 CCACCTCAGCCTCTAGCATATAAACTAGACACATCCTCATCTTTTGAGGTCTAATCATTT 163437
QY 61 GGATTTTGTCTTTCAGATGGCTTTCCCGAAGATGAGATTGATGTATATATCTGCCCTTCG 120
DB 163438 GGATTTTGTCTTTCAGATGGCTTTCCCGAAGATGAGATTGATGTATATATTTGCCTTCG 163497
QY 121 GTTCTGGGGGCTCTTTGTTGTATTTTAGCATGTACAGTCTAAATCTCTTCAAAGAACAG 180
DB 163498 GTTCTGGGGGCTCTTTGTTGTATTTTAGCATGTACAGTCTAAATCTCTTCAAAGAACAG 163557
QY 181 TCCTTTGTTTACAAAGAACGCGGAATTCCTTAAGCTCCAGATACAGACTGCAGGGCAG 240
DB 163558 TCCTTTGTTTACAAAGAACGCGGAATTCCTTAAGCTCCAGATACAGACTGCAGGGCAG 163617
QY 241 ACACCTCCCTTCCTCGCTCGTCTGTCACCTCATCCCAAAACAGTTGGCTGAGCGCATG 300
DB 163618 ACACCTCCCTTCCTCGCTCGTCTGTCACCTCATCCCAAAACAGTTGGCTGAGCGCATG 163677
QY 301 GCCATCCGGCAGACGTGGGGGAAAAGAGAGGACGTTGAAGGGAAGACAGCTGAAGACATTC 360
DB 163678 GCCATCCGGCAGACGTGGGGGAAAAGAGAGGATGTTGAAGGGAAGACAGCTGAAGACATTC 163737
QY 361 TTCTCTCTGGGGACCCAGCCAGTGCAGCGGAAACAAAGAGGTGGACAGAGAGCCAG 420
DB 163738 TTCTCTCTGGGGACCCAGCCAGTGCAGCGGAAACAAAGAGGTGGACAGAGAGCCAG 163797
QY 421 CGACA CGGGGACATTTATCCAGAAGGATTTCTAGACGTTATTACAATCTGACCTCGAAG 480
DB 163798 CGACA CGGGGACATTTATCCAGAAGGATTTCTAGACGTTATTACAATCTGACCTCGAAG 163857
QY 481 ACCATGATGGGCATAGAAATGGGTCCATCGCTTTTGTCTCAGGGCGCGTTTGTGATGAAA 540
DB 163858 ACCATGATGGGCATAGAAATGGGTCCATCGCTTTTGTCTCAGGGCGCGTTTGTGATGAAA 163917
QY 541 ACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGAATCTCTCAAGGAAGAAC 600
DB 163918 ACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGAATCTCTCAAGGAAGAAC 163977
QY 601 AGAACAAACAGGTTTTTTCACCTGGCTTTCTTGAATCTCAATGAGTTTCCCATCAGGCGCA 660
DB 163978 AGAACAAACAGGTTTTTTCACCTGGCTTTCTTGAATCTCAATGAGTTTCCCATCAGGCGCA 164037
QY 661 TTACAGCAAGTGGTTTGTTCAGTAAATCTGAATATCCGTGGGACAGGTACCCACCAATTCG 720
DB 164038 TTACAGCAAGTGGTTTGTTCAGTAAATCTGAATATCCGTGGGACAGGTACCCACCAATTCG 164097
QY 721 TCCGGCACCGCTACGTGTTTCTGGCGACGTGGCGAGTCAAGTGTACAATGTCTCCAG 780
DB 164098 TCCGGCACCGCTACGTGTTTCTGGCGACGTGGCGAGTCAAGTGTACAATGTCTCCAG 164157
QY 781 AGCGTCCCATACATTAATACTGGAAGAGCTGTTTCTGGGGCTCTGCCTCGAAGGCTGAAC 840
DB 164158 AGCGTCCCATACATTAATACTGGAAGAGCTGTTTCTGGGGCTCTGCCTCGAAGGCTGAAC 164217
QY 841 ATCAGATTGGAGGAGCTCCACTCCAGCCGACCTTTTTCAGGGGGGTTTACGCTTCTCC 900
DB 164217 ATCAGATTGGAGGAGCTCCACTCCAGCCGACCTTTTTCAGGGGGGTTTACGCTTCTCC
```



```
QY 901 GTATGCTCTTTCAGAGGATCGTGCCTGCCACTTCATCAAGCTCGGACTCTCTTGGAC 960
Db 87493 GTATGCTCTTTCAGAGGATCGTGCCTGCCACTTCATCAAGCTCGGACTCTCTTGGAC 87552
QY 961 TACTGGCAGGCTCTAGAGAAATCCCGGGGGGAGAGATTGTCGCCCTGTCTGA 1011
Db 87553 TACTGGCAGGCTCTAGAGAAATCCCGGGGGGAGAGATTGTCGCCCTGTCTGA 87603

RESULT 6
AB041416 1576 bp DNA linear PRI 17-OCT-2000
LOCUS Homo sapiens beta1,3-GalT 5 gene for UDP-Gal:GlcNAc
DEFINITION beta1,3-galactosyltransferase 5, partial cds.
ACCESSION AB041416
VERSION AB041416.1 GI:7593026
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1576)
AUTHORS Liu, Y. and Saitou, N.
TITLE Silver Project
JOURNAL Published Only in DataBase (2000)
REFERENCE 2 (bases 1 to 1576)
AUTHORS Liu, Y. and Saitou, N.
TITLE Direct Submission
JOURNAL Submitted (11-APR-2000) Naruya Saitou, National Institute of
Genetics, Laboratory of Evolutionary Genetics; 1111 Yata, Mishima,
Shizuoka 411-8540, Japan (E-mail:nsaitou@genes.nig.ac.jp,
URL: http://saver.lab.nig.ac.jp/~silver/, Tel: 81-559-81-6790,
Fax: 81-559-81-6789)
FEATURES
Location/Qualifiers
1..1576
/organism="Homo sapiens"
/mol_type="genomic DNA"
/isolate="#056"
/db_xref="taxon:9606"
/note="human sequence used for primer design based on Acc#
AB020337"
<1..27
/exon
/number=3
intron
28..674
/number=3
gene
675..1576
/gene="beta1,3-GalT 5"
675..1576
/gene="beta1,3-GalT 5"
/codon_start=1
/product="UDP-Gal:GlcNAc beta1,3-galactosyltransferase 5"
/protein_id="BAA94501.1"
/db_xref="GI:7593027"
/translation="MAPEPMRLMYICLVLALGALCLYFSMYSINLPFKESFYVKDGNF
LKLPDTCRQTPPELVLLVTSRHLQALRMALRQVTKGKQLKLPFLGLGTSS
AAETKEVDQSORHGDIIQKDFLDVYNNLTLMWLEWVHRFCPOAFAVTKTQSDMF
INVDYLLKLRNTRFPFGLKNEFPPIQPPKSFVSKSYPRDPRYPFPCSGTG
YVFGSDVASQYINVSQVPIYKLDVFGVLGCLERLRLBELHSQTPFPFGGLRFSVC
LFRIRVACHFKPTLLDYQWALE"
675..1576
/exon
/gene="beta1,3-GalT 5"
/number=4

ORIGIN
Query Match 96.8%; Score 978.4; DB 9; Length 1576;
Best Local Similarity 99.9%; Pred. No. 9.4e-261;
Matches 979; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCACCTAGCCTCTAGCATAAACTAGACATCCTCATGCTTTGAGGTCTAATCAT 60
Db 597 CCACCTAGCCTCTAGCATAAACTAGACATCCTCATGCTTTGAGGTCTAATCAT 656
```

```
QY 61 GGATTTTGTTCCTTTTCAGATGGCTTTCCGAGATGAGATTGATATATATCTGCCTTCTG 120
Db 657 GGATTTTGTTCCTTTTCAGATGGCTTTCCGAGATGAGATTGATATATATTTGGCTTCTG 716
QY 121 GTTCTGGGGGCTCTTTGTTGTTATTTTAGCATGTACAGTCTTAAATCCCTTTCAAGAACAG 180
Db 717 GTTCTGGGGGCTCTTTGTTGTTATTTTAGCATGTACAGTCTTAAATCCCTTTCAAGAACAG 776
QY 181 TCCCTTTGTTTACAAGAAAGACGGGAACTTCTTAAGCTCCAGATACAGACTCGAGGGCAG 240
Db 777 TCCCTTTGTTTACAAGAAAGACGGGAACTTCTTAAGCTCCAGATACAGACTCGAGGGCAG 836
QY 241 ACACCTCCCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
Db 837 ACACCTCCCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 896
QY 301 GCCATCCGGCAGAGCTGGGGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
Db 897 GCCATCCGGCAGAGCTGGGGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 956
QY 361 TTCTCTCTGGGGGACACCAGCAGTGCAGCGGAAACAAAGAGAGGTGGACAGGAGAGCCAG 420
Db 957 TTCTCTCTGGGGGACACCAGCAGTGCAGCGGAAACAAAGAGAGGTGGACAGGAGAGCCAG 1016
QY 421 CGACACGGGGACATTTATCCAGAGAGATTTCTAGACGTCTATTACAATCTGACCTGAAG 480
Db 1017 CGACACGGGGACATTTATCCAGAGAGATTTCTAGACGTCTATTACAATCTGACCTGAAG 1076
QY 481 ACCATGATGGGCATAGATGGTCCATCGCTTTTCTCTCAGGGGGCTGTTGTGATGAAA 540
Db 1077 ACCATGATGGGCATAGATGGTCCATCGCTTTTCTCTCAGGGGGCTGTTGTGATGAAA 1136
QY 541 ACAGACTCAGACATGTTTCATCAATGTTGACATCTGACTGAACTGCTTCTCAAGAAAAAC 600
Db 1137 ACAGACTCAGACATGTTTCATCAATGTTGACATCTGACTGAACTGCTTCTCAAGAAAAAC 1196
QY 601 AGAACACACAGGTTTTTTCATCTGGCTTTCTTGAACCTCAATGAGTTTCCATCAGGAGCCA 660
Db 1197 AGAACACACAGGTTTTTTCATCTGGCTTTCTTGAACCTCAATGAGTTTCCATCAGGAGCCA 1256
QY 661 TTCAGCAAGTGTGTCAGTAAATCTGAATATCGTGGGACAGGTACCCACCATCTCTGC 720
Db 1257 TTCAGCAAGTGTGTCAGTAAATCTGAATATCGTGGGACAGGTACCCACCATCTCTGC 1316
QY 721 TCCGGCACCGGCTACGTGTTTCTGGGACGTGCGAGTCAAGTGTCAATGTTCTCTCAAG 780
Db 1317 TCCGGCACCGGCTACGTGTTTCTGGGACGTGCGAGTCAAGTGTCAATGTTCTCTCAAG 1376
QY 781 AGCGTCCCATACATTAACCTGGAAGAGCGTGTGTTGGGGCTCTGCTCGAAAGGCTGAAC 840
Db 1377 AGCGTCCCATACATTAACCTGGAAGAGCGTGTGTTGGGGCTCTGCTCGAAAGGCTGAAC 1436
QY 841 ATCAGATTGAGGAGCTCCATCCAGCGGACCTTTTTCAGGGGGCTTAGCTTCTCC 900
Db 1437 ATCAGATTGAGGAGCTCCATCCAGCGGACCTTTTTCAGGGGGCTTAGCTTCTCC 1496
QY 901 GTATGCTCTTTCAGAGGAGTCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 960
Db 1497 GTATGCTCTTTCAGAGGAGTCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 1556
QY 961 TACTGGCAGGCTCTAGAGAA 980
Db 1557 TACTGGCAGGCTCTAGAGAA 1576

RESULT 7
AB041415 1570 bp DNA linear PRI 13-APR-2000
LOCUS Pan paniscus beta1,3-GalT 5 gene for UDP-Gal:GlcNAc
DEFINITION beta1,3-galactosyltransferase 5, partial cds.
ACCESSION AB041415
VERSION AB041415.1 GI:7593024
KEYWORDS
```


Qy	601	AGAACACACAGGTTTTTCACTGGCTTCTTGAAACTCAATGAGTTTCCCATCAGCGCAGCCA	660
Db	1197	AGAACACACAGGTTTTTCACTGGCTTCTTGAAACTCAATGAGTTTCCCATCAGCGCAGCCA	1256
Qy	661	TTCAGCAAGTGGTTTTGTCTAGTAATATCCGTGGGACAGGTACCCACCATTTCTGC	720
Db	1257	TTCAGCAAGTGGTTTTGTCTAGTAATATCCGTGGGACAGGTACCCACCATTTCTGC	1316
Qy	721	TCCGGCACCGGTACGTGTTTTCTGCGACGTCGGCAGTCAGGTGTACAATGTCTCCAAG	780
Db	1317	TCCGGCACCGGTACGTGTTTTCTGCGACGTCGGCAGTCAGGTGTACAATGTCTCCAAG	1376
Qy	781	AGCGTCCCATACATTTAACTGGGAAGCGTGTGTGGGGCTCTGCCTCGAAGAGGCTGGAAC	840
Db	1377	AGCGTCCCATACATTTAACTGGGAAGCGTGTGTGGGGCTCTGCCTCGAAGAGGCTGGAAC	1436
Qy	841	ATCAGATTGGAGGAGTCCACTCCACGCGACCTTTTTTCCAGGGGCTTTACGCTTCTCC	900
Db	1437	ATCAGATTGGAGGAGTCCACTCCAGCGGACCTTTTTTCCAGGGGCTTTACGCTTCTCC	1496
Qy	901	GTATGCTCTTTCAGGAGGATCGTGGCGTCGCCATTCATCAAGCCTCGGACTCTTTGGAC	960
Db	1497	GTATGCTCTTTCAGGAGGATCGTGGCGTCGCCATTCATCAAGCCTCGGACTCTC-TGGAC	1555
Qy	961	TACTGGCAGG	970
Db	1556	TACTGGCAGG	1565

1	Qy	CGACCTCAGGCTCTCTAGCATAAACCTAGACACATCCTCATGCTTTTGAGGTCCTAAATCATTT	60
597	Db	CCACCTCAGGCTCTCTAGCATAAACCTAGACACATCCTCATGCTTTTGAGGTCCTAAATCATTT	656
61	Qy	GGATTTGTTCTCTTTCAGATGGCTTTTCCCAAGATGAGATTGATGTATATCTGCCCTTCG	120
657	Db	GGATTTGTTCTCTTTCAGATGGCTTTTCCCAAGATGAGATTGATGTATATCTGCCCTTCG	716
121	Qy	GTTCTGGGGCTCTTTGTTGTTGATATTTTAGCATGTACAGTCTAAATCCTTTCAAAGAACAG	180
717	Db	GTTCTGGGGCTCTTTGTTGTTGATATTTTAGCATGTACAGTCTAAATCCTTTCAAAGAACAG	776
181	Qy	TCCTTTGTTTACAAGAAAGACGGGAATTCCTTAAGCTCCAGATACAGACTCGAGGCAG	240
777	Db	TCCTTTGTTTACAAGAAAGACGGGAATTCCTTTAAGCTCCAGATACAGACTCGAGGCAG	836
241	Qy	ACACCTCCCTTCCTCGTCTGCTGGTGACCTCATCCCACAAACAGTTGGCTGAGCGCATG	300
837	Db	ACACCTCCCTTCCTCGTCTGCTGGTGACCTCATCCCACAAACAGTTGGCTGAGCGCATG	896
301	Qy	GCCATCCGGCAGACGTGGGGGAAAGAGAGACCGTGAAAGGAAAGCAGCTGAAAGACATTC	360
897	Db	GCCATCCGGCAGACGTGGGGGAAAGAGAGACCGTGAAAGGAAAGCAGCTGAAAGACATTC	956
361	Qy	TTCTCTCTGGGACACACAGCAGTGCAGCGGAAACAAAGAGGTGGACACAGAGAGCCAG	420
957	Db	TTCTCTCTGGGACACACAGCAGTGCAGCGGAAACRAAAGAGGTGGACACAGAGAGCCAG	1016
421	Qy	CGACAGGGGACATTTACCAAGAGGATTTCTAGACGTCTATTACAATCTGACCCCTGAAG	480
1017	Db	CGACAGGGGACATTTACCAAGAGGATTTCTAGACGTCTATTACAATCTGACCCCTGAAG	1076
481	Qy	ACCATGATGGGCATAGAAATGGGTCCTATCGCTTTTGTCTCAGCGCGCGTTTGTGATGAAA	540
1077	Db	ACCATGATGGGCATAGAAATGGGTCCTATCGCTTTTGTCTCAGCGCGCGTTTGTGATGAAA	1136
541	Qy	ACAGACTCAGACATGTTTCATCAATGTTTGAATCTATCTGACTGAATCTGTTCTGAAGAAAAAC	600
1137	Db	ACAGACTCAGACATGTTTCATCAATGTTTGAATCTATCTGACTGAATCTGTTCTGAAGAAAAAC	1196

```

RESULT 9
AB041412 linear PRI 17-OCT-2000
LOCUS
DEFINITION
Gorilla gorilla betal,3-Galt 5 gene for UDP-Gal:GlcNAc
betal,3-galactosyltransferase 5, partial cds.
ACCESSION
AB041412
VERSION
AB041412.1 GI:7593018
KEYWORDS
Gorilla gorilla (gorilla)
SOURCE
Gorilla gorilla
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
REFERENCE
1 (bases 1 to 1360)
AUTHORS
Liu, Y. and Saitou, N.
TITLE
Silver Project
JOURNAL
Published only in DataBase (2000)
REFERENCE
2 (bases 1 to 1360)
AUTHORS
Liu, Y. and Saitou, N.
TITLE
Direct Submission
JOURNAL
Submitted (11-APR-2000) Naryu Saitou, National Institute of
Genetics, Laboratory of Evolutionary Genetics; 1111 Yata, Mishima,
Shizuoka 411-8540, Japan (E-mail:nsaitou@genes.nig.ac.jp,
URL:http://saver.lab.nig.ac.jp/~silver/, Tel:81-559-81-6790,
Fax:81-559-81-6789)
FEATURES
Location/Qualifiers
1..1360
/organism="Gorilla gorilla"
/mol_type="genomic DNA"
/isolate="#085"
/db_xref="taxon:9593"
/notes="human sequence used for primer design based on Acc#
AB020337"
<1..466
/number=3
467..1360
/gene="betal,3-Galt 5"
467..>1360
/gene="betal,3-Galt 5"
/codon_start=1
/product="UDP-Gal:GlcNAc betal,3-galactosyltransferase 5"
/protein_id="BAA9497.1"
/db_xref="GI:7593019"
/translat="MAFFPKRLMYICLLVIGALCYFMSYSLNPFKEQSFVYKKQVNF
IKIPDRCOTPEPVLIVTSSHKALAEAMATROTWKERTVKGKOLKTFELIGTSS
intron
gene
CDS

```

exon

Query Match	93.8%	Score 948.4	DB 9	Length 1360
Best Local Similarity	98.3%	Pred. No. 2.1e-252		
Matches 955	Conservative 3	Mismatches 14	Indels 0	Gaps 0
QY	1	CCACCTCAGGCTCTCTAGCATATAAACTAGACACATCCCTCATGCTTTTCAGGTCTTAATCATTT	60	
DB	389	CCACCTCAGGCTCTCTAGCATATAAACTAGACACATCCCTCATGCTTTTCAGGTCTTAATCATTT	448	
QY	61	GGATTTTGTTCCTTTTCAGATGGCTTTTCCCGAAGATGAGATTGATGTATATCTGCCTTCTG	120	
DB	449	GGATTTTGTTCCTTTTCAGATGGCTTTTCCCGAAGATGAGATTGATGTATATTTGCCTTCTG	508	
QY	121	GTTCTGGGGCTCTTTGTTTGTATTTAGCATGTACAGTCTAAATCTCTTCAAAGAAACAG	180	
DB	509	GTTCTGGGGCTCTTTGTTTGTATTTAGCATGTACAGTCTGAATCTCTTCAAAGAAACAG	568	
QY	181	TCCTTTGTTTACAGAAAGACGGGAACCTTCCCTTAAGCTCCACAGATACAGACTCGAGGCG	240	
DB	569	TCCTTTGTTTACAGAAAGACGTGAACCTTCTTAAAGCTCCAGATACAGACTCGAGGCG	628	
QY	241	ACACCTCCCTTCCTCGTCTGCTGGTGAACCTCATCCCAAAACAGTTGGCTGAGCGCATG	300	
DB	629	ACACCTCCCTTCCTCGTCTGCTGGTGAACCTCATCCCAAAACAGTTGGCTGAGCGCATG	688	
QY	301	GCAATCGGCAGACGTGGGGAAAGACAGACAGACGCTGAAGGAAACAGCTGAAGACATTC	360	
DB	689	GCCATCAGGCAGACGTGGGGAAAGACAGACAGTGAAGGAAACAGCTGAAGACATTC	748	
QY	361	TTCTCTCTGGGGACCCACACAGCTGCAGCGGAAACAAAGAGGTGGACACGAGGAGGCCAG	420	
DB	749	TTCTCTCTGGGGACCCACAGCGTGCAGCGGAAACAAAGAGGTGGACACGAGGAGGCCAG	808	
QY	421	CGACACGGGACATTTATCCAGAAGATTTCTCTAGACGCTCTATTAACAATCTGACCCCTGAAG	480	
DB	809	CGACACGGGACATTTATCCAGAAGATTTCTCTAGAYGCTCTATTACATCTGACCCCTGAAG	868	
QY	481	ACCATGATGGGCATAGAAATGGGTCATTCGCTTTTGTCTCAGCGCGGCTTTGTGATGAAA	540	
DB	869	ACCATGATGGGCATAGAAATGGGTCATTCGCTTTTGTCTCAGCGCGGCTTTGTGATGAAA	928	
QY	541	ACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGAATCTGCTTCTGAAGAAAAC	600	
DB	929	ACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGAATCTGCTTCTGAAGAAAAC	988	
QY	601	AGAACACACAGGTTTTTTCACCTGGCTTTCTTGAACCTCAATGAGTTTCCCATCAGGACGCCA	660	
DB	989	AGAACACACAGGTTTTTTCACCTGGCTTTCTTGAACCTCAATGAGTTTCCCATCAGGACGCCA	1048	
QY	661	TTGACGAAGTGGTTTGTTCAGTAAATCTGAATATCCGTGGGACAGGTATCCCAATCTGTC	720	
DB	1049	TTGACGAAGTGGTTTGTTCAGTAAATCTGAATATCCGTGGGACAGGTATCCCAATCTGTC	1108	
QY	721	TCCGGCACCGGCTACGTTGTTTTCTGGCGAGCTGGCGAGTCAAGTGTACATGTCTCCAAG	780	
DB	1109	TCCGGCACCGGCTACGTTGTTTTCTGGCGAGCTGGCGAGTCAAGTGTACATGTCTCCAAG	1168	
QY	781	AGCGTCCCATATCATTTAACTGGAAGACGTGTTTGTGGGGCTCTGCTCTGAAAGGCTCAAC	840	
DB	1169	AGCGTCCCATATCATTTAACTGGAAGACGTGTTTGTGGGGCTCTGCTCTGAAAGGCTCAAC	1228	
QY	841	ATCAGATTGAGGAGCTCCATCTCCAGCCGACCTTTTTTTCAGGGGGCTTACGCTTCTCC	900	
DB	1229	ATCAGATTGAGGAGCTCCATCTCCAGCCGACSTTTTTTCCAGGGGGCTTACGCTTCTCY	1288	

Qy 901 GTATGCTCTTCAGGAGGATCGTGGCTCGCACTTATCAAGCCTCGGACTCTCTTGGAC 960
 Db 1289 GTATGCGCTTCAGGAGGATCGTGGCTCGCACTTATCAAGCCTCGGACTCTCTTGGAC 1348
 Qy 961 TACTGGCAGGCT 972
 Db 1349 TACTGGCAGGCT 1360

RESULT 10
 AB041414 1566 bp DNA linear PRI 13-APR-2000
 DEFINITION Pan troglodytes betal.3-galt 5 gene for UDP-Gal:GlcNAc
 betal.3-galactosyltransferase 5, partial cds.
 ACCESSION AB041414
 VERSION
 KEYWORDS
 SOURCE Pan troglodytes (chimpanzee)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
 REFERENCE 1 (bases 1 to 1566)
 AUTHORS Liu, Y. and Saitou, N.
 TITLE Silver Project
 JOURNAL Published Only in DataBase (2000)
 REFERENCE 2 (bases 1 to 1566)
 AUTHORS Liu, Y. and Saitou, N.
 TITLE Direct Submission
 JOURNAL Submitted (11-APR-2000) Naruya Saitou, National Institute of
 Genetics, Laboratory of Evolutionary Genetics, 111 Yata, Mishima,
 Shizuoka 411-8540, Japan (E-mail:nsaitou@genes.nig.ac.jp,
 URL: http://sayer.lab.nig.ac.jp/~silver/, Tel:81-559-81-6790,
 Fax:81-559-81-6789)

FEATURES
 Location/Qualifiers
 1..1566
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /isolate="chimp-202"
 /db_xref="taxon:9598"
 /note="human sequence used for primer design based on Acc#
 AB020337"
 exon <1..28
 intron 29..675
 gene 675..1566
 /gene="betal.3-Galt 5"
 CDS 675..>1566
 /gene="betal.3-Galt 5"
 /codon_start=1
 /product="UDP-Gal:GlcNAc betal.3-galactosyltransferase 5"
 /protein_id="BA094499.1"
 /db_xref="GI:7593023"
 /translation="MAFPKRLMYCVLLVGLCVYFSMYSLNLFKEQSFVYKDGNF
 LKLPDCRTPPFPVLLVTSRRQALAEIRMTGWKERTVKGKQLKTFLLGTSSD
 AATKEVDQSSQSDGDIIOKDFLDVYVNLTKTMWGLEWHRFCPOAAFPVKMTDSDF
 INVDYLTLLKKNRTRPFTGKLNKEFPIROFSPKWFVSKSEYKPDVRYPPFCSGFG
 YVFSQVASGVYNNVSSVPYKLEDFVGLCLERLNIRLEELHSQTPFFPGGLRFSVCS
 YFFRIVACHFIPKPTLDYVQ"

exon 675..>1566
 /gene="betal.3-Galt 5"
 /number=4

ORIGIN

Query Match 93.5%; Score 945.6; DB 9; Length 1566;
 Best Local Similarity 98.4%; Pred. No. 1.3e-251;
 Matches 954; Conservative 1; Mismatches 15; Indels 0; Gaps 0;

Qy 1 CCACCTCAGCTCTTAGCATAAACTAGACACATCCTCATGCTTTTCAGGCTTAATCATT 60
 Db 597 CCACCTCAGCTCTTAGCATAAACTAGACACATCCTCATGCTTTTCAGGCTTAATCATT 656
 Oy 61 GGATTTTGTTCCTTTCAGATGGCTTTTCCCGAAGATGAGATTGATGTATATCTGCCTTCTG 120


```
Db 944 TCCACTCCAGCGACCTTTTTCAGGGGCTTACGCTTCTCCGTATGCTCTTCAGGA 1003
Qy 917 GGATCGTGGCTGCCACTTCAATCAAGCCCTCGAGCTCTCTTGGACTACTTGGCAGGCTCTAG 976
Db 1004 GGATCGTGGCTGCCACTTCAATCAAGCCCTCGAGCTCTCTTGGACTACTTGGCAGGCTCTAG 1063
Qy 977 AGAATTCCCGGGGGAAGATTGTCGCGCTGCTGA 1011
Db 1064 AGAATTCCCGGGGGAAGATTGTCGCGCTGCTGA 1098

RESULT 12
HSA6078 HSA6078 933 bp DNA linear PRI 11-MAY-2000
LOCUS Homo sapiens beta3gal-T6 gene.
DEFINITION
ACCESSION AJ006078
VERSION beta-1.3-galactosyltransferase; beta3gal-T6 gene.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Amado,M., Carneiro,F. and Clausen,H.
TITLE Cloning and expression of two beta-1,3-galactosyltransferases:
beta3gal-T5 and beta3gal-T6
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 933)
AUTHORS Amado,M.
TITLE Direct Submission
JOURNAL Submitted (11-MAY-1998) Amado M., Department of Oral Diagnostics,
Royal dental School, Norre Alle 20, 2200 Copenhagen, DENMARK
FEATURES
source
1. .933
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
1. .933
/gene="beta3gal-T6"
1. .933
/gene="beta3gal-T6"
/codon_start=1
/product="beta-1,3-galactosyltransferase"
/protein_id="CAB91547.1"
/db_xref="GI:7799923"
/db_xref="GOA:Q9Y2C3"
/db_xref="UniProt/Swiss-Prot:Q9Y2C3"
/translation="MAPPMRLMVICLLVLGLCLXYSMNSLNPFPKESFYVKDGNF
LKLPTDCROTPPELVLLVTSKQLAERMAIROTGWKERTVKQKLTPELLGTTSS
AAETKEVDQESQRHGDIIQKDLVDVYNLTILKTMGIEWHRFCQAAPVMTSDMF
INDVLTLLKRNKTRTFYFKLINEFFIRQPPKMFVSKSBYPMDRYPFFCSGTG
YVFGSDVASQVYNSKVPYIKJEDVFVGLCLERLNLRLLELSQPTFFPPGGLRFSVC
LFRRIVACHFIKPRTLDDYMQALENSRGEDCPVPV"

ORIGIN
Query Match 91.9%; Score 928.8; DB 9; Length 933;
Best Local Similarity 99.7%; Pred. No. 5.9e-24;
Matches 930; Conservative 0; Mismatches 37; Indels 0; Gaps 0;
Qy 79 ATGGCTTCCCGAAGATGAGATTGATATATCGCCCTTCTGGTCTCGGGGCTCTTTGT 138
Db 1 ATGGCTTCCCGAAGATGAGATTGATATATCGCCCTTCTGGTCTCGGGGCTCTTTGT 60
Qy 139 TTGTATTTTACATGATACAGTCTAAATCCTTTCAAAGAACAGTCTTTGTTTACAAGAAA 198
Db 61 TTGTATTATAGCATGAACAGTCTAAATCCTTTCAAAGAACAGTCTTTGTTTACAAGAAA 120
Qy 199 GACGGGAATCTCTTAAGCTCCAGATACAGACTGACGAGGAGACACCTCCCTTCCTGTC 258
Db 121 GACGGGAATCTCTTAAGCTCCAGATACAGACTGACGAGGAGACACCTCCCTTCCTGTC 180
Qy 259 CTGCTGGTGACCTCATCCCAACACAGTTGGTGTAGGCGCATCGGCCATCGCGCAGCGTGG 318
```

```
Db 181 CTGCTGGTGACTCATCCCAAAACAGTTGGCTGAGCGCATGGCCATCCGGCAGAGCTGG 240
Qy 319 GGGAAAGAGAGAGCGGTGAAGGGAAAGCAGCTGAAGACATTTCTTCTCTGGGACCAACC 378
Db 241 GGGAAAGAGAGAGCGGTGAAGGGAAAGCAGCTGAAGACATTTCTTCTCTGGGACCAACC 300
Qy 379 AGCAGTCGACGGGAAAACAAAGAGGTGGACGAGGAGCCAGCGACACGGGGACATTATC 438
Db 301 AGCAGTCGACGGGAAAACAAAGAGGTGGACGAGGAGCCAGCGACACGGGGACATTATC 360
Qy 439 CAGAAGGATTTCTTAGACGCTATTATAATCTGACCCCTGAAGACCATGATGGGCATAGAA 498
Db 361 CAGAAGGATTTCTTAGACGCTATTATAATCTGACCCCTGAAGACCATGATGGGCATAGAA 420
Qy 499 TGGGTCCATCGCTTTTGTCTCAGCGCGGTTTGTGATGAAAAACAGACTCAGACATGTTT 558
Db 421 TGGGTCCATCGCTTTTGTCTCAGCGCGGTTTGTGATGAAAAACAGACTCAGACATGTTT 480
Qy 559 ATCAATGTTGACTATCTGACTGAACTGCTCTGAAGAAAACAGAAACAGAGGTTTTTC 618
Db 481 ATCAATGTTGACTATCTGACTGAACTGCTCTGAAGAAAACAGAAACAGAGGTTTTTC 540
Qy 619 ACTGGCTTTTGAACCTCAATGAGTTTCCCATCAGGCGAGCCATTCAGCAAGTGGTTTCTC 678
Db 541 ACTGGCTTTTGAACCTCAATGAGTTTCCCATCAGGCGAGCCATTCAGCAAGTGGTTTCTC 600
Qy 679 AGTAAATCTGAATATCCGTGGGACAGGTACCCACATTTCTCTCGGACACGGGTACGTG 738
Db 601 AGTAAATCTGAATATCCGTGGGACAGGTACCCACATTTCTCTCGGACACGGGTACGTG 660
Qy 739 TTTTCTGGCGAGCTGGGAGTCTGAGTGTACAACTCTCCAGAGCGTCCCATACATTAAA 798
Db 661 TTTTCTGGCGAGCTGGGAGTCTGAGTGTACAACTCTCCAGAGCGTCCCATACATTAAA 720
Qy 799 CTGAAGACGCTGTTTGTGGGGCTCTGCCTCGAAAGGCTGAACATCAGATTGGAGGAGCTC 858
Db 721 CTGAAGACGCTGTTTGTGGGGCTCTGCCTCGAAAGGCTGAACATCAGATTGGAGGAGCTC 780
Qy 859 CACTCCACGCGACCTTTTTCAGGGGGCTTACGCTTCTCCGTATGCTCTTTCAGGAGG 918
Db 781 CACTCCACGCGACCTTTTTCAGGGGGCTTACGCTTCTCCGTATGCTCTTTCAGGAGG 840
Qy 919 ATCGTGGCTGCCACTTCATCAAGCCTCGGACTCTTGGACTACTGCGAGGCTCTAGAG 978
Db 841 ATCGTGGCTGCCACTTCATCAAGCCTCGGACTCTTGGACTACTGCGAGGCTCTAGAG 900
Qy 979 AATTCCCGGGGGAAGATTGTCCGCTGCTCTGA 1011
Db 901 AATTCCCGGGGGAAGATTGTCCGCTGCTCTGA 933

RESULT 13
AB020337 Homo sapiens mRNA for UDP-Gal:GlcNAc beta1,3-galactosyltransferase
LOCUS 5, complete cds.
DEFINITION
ACCESSION AB020337
VERSION AB020337.1 GI:4835502
KEYWORDS UDP-Gal:GlcNAc beta1,3-galactosyltransferase 5.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (sites)
AUTHORS Ieshiki,S., Togayachi,A., Kudo,T., Nishihara,S., Watanabe,M.,
Kubota,T., Kitajima,M., Shiraishi,N., Sasaki,K., Andoh,T. and
Narimatsu,H.
TITLE Cloning, expression, and characterization of a novel
UDP-galactose:beta-N-acetylglucosamine
beta1,3-galactosyltransferase (beta3Gal-T5) responsible for
synthesis of type 1 chain in colorectal and pancreatic epithelia
and tumor cells derived therefrom
```

J. Biol. Chem. 274 (18), 12499-12507 (1999)

99230269

10212226

(bases 1 to 2762)

Ishiki, S., Togayachi, A. and Narimatsu, H.

Direct Submission

Submitted (20-NOV-1998) Hisashi Narimatsu, Soka University,

Institute of Life Science; 1-236, Tangi-cho, Hachioji, Tokyo

192-8577, Japan (E-mail: sieshiki@po.ijinet.or.jp,

Tel: 81-426-91-9466, Fax: 81-426-91-9315)

Location/Qualifiers

1..2762

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/chromosome="21"

/map="21q22.3"

/cell_lines="Colo 205"

/cell_type="Adenocarcinoma"

1..2762

/gene="betal,3-Galt 5"

1..273

/gene="betal,3-Galt 5"

/number=1

274..433

/gene="betal,3-Galt 5"

/number=3

434..2762

/gene="betal,3-Galt 5"

/number=4

434..1366

/gene="betal,3-Galt 5"

/codon_start=1

/product="UDP-Gal:GlcNAc betal,3-galactosyltransferase 5"

/protein_id="BAA77664.1"

/db_xref="GI:4835503"

/translation="MAPEKRLMYICLLVLCALCLYFSMSLNPFKQSFVYKDGNF
LKLPDTCRQPPFLVLLVTSVSHKQLAEMAIQVWGERVWVKGQKLTFFLLGTSS
AAETKEVDQSGHDIQKDFLDVYVNLTKTWGIEWVHRVPCQAAVFNKDTSDMF
INVLYTELLAKNLTFRFFGFLKLINEFPIROPSPKWFVSKSEVPMDRYPPFCSTG
YVFGDVASQVYNYVSKSPYIKLEDFVFGCLERLRLNLESLHSQTPFPFGLRPSVC
LFRIVACHFIKPTLLDYQALENSRGEDCPV"

polyA_site

/gene="betal,3-Galt 5"

/note="45 a nucleotides"

ORIGIN

Query Match 91.8%; Score 928.6; DB 9; Length 2762;

Best Local Similarity 99.6%; Pred. No. 7e-247;

Matches 931; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 77 AGATGGCTTCCGAGATGAGATTGATATCTGCTTCTGCTGCTGGGGCTCTTT 136

DB 432 AAATGGCTTCCGAGATGAGATTGATATTTGCTTCTGCTGCTGGGGCTCTTT 491

QY 137 GTTTGATTTTAGCATGTACAGTCTAAATCCTTTCAAAGACAGTCCTTTGTGTACAGA 196

DB 492 GTTTGATTTTAGCATGTACAGTCTAAATCCTTTCAAAGACAGTCCTTTGTGTACAGA 551

QY 197 AAGACGGGAACCTCTTAAGCTCCAGATACAGACTGCAGGCAGACACCTCCCTTCCTCG 256

DB 552 AAGACGGGAACCTCTTAAGCTCCAGATACAGACTGCAGGCAGACACCTCCCTTCCTCG 611

QY 257 TCCTGCTGTGACCTCATCCCAACACAGTTGGCTGAGCGCATCGGCAGACGT 316

DB 612 TCCTGCTGTGACCTCATCCCAACACAGTTGGCTGAGCGCATCGGCAGACGT 671

QY 317 GGGGAAAGAGAGACGCGTGAAGGAAAGACAGCTGAAGACATTTCTCTCTGGGGACCA 376

DB 672 GGGGAAAGAGAGAGTGTGAAGGAAAGACAGCTGAAGACATTTCTCTCTGGGGACCA 731

QY 377 CCACAGTGCAGCGGAACAAAGAGGTGGACAGGAGCCAGCGACACCGGGGACATTA 436

DB 1331 CCACAGTGCAGCGGAACAAAGAGGTGGACAGGAGCCAGCGACACCGGGGACATTA 791

QY 437 TCCAGAAAGGATTTCTAGACGTCTATTACATCTGACCTCGAACACCATGATGGCATAG 496

DB 792 TCCAGAAAGGATTTCTAGACGTCTATTACATCTGACCTCGAACACCATGATGGCATAG 851

QY 497 AATGGGTCCATCGCTTTTGTCTCAGCGGGGCTTTGTGATGAAACACAGCTCAGACATGT 556

DB 852 AATGGGTCCATCGCTTTTGTCTCAGCGGGGCTTTGTGATGAAACACAGCTCAGACATGT 911

QY 557 TCATCAATGTTGACTATCTGACTCACTGCTTCTGAGAAAAACAGAACACAGGTTTTT 616

DB 912 TCATCAATGTTGACTATCTGACTCACTGCTTCTGAGAAAAACAGAACACAGGTTTTT 971

QY 617 TCACTGGCTTCTTGAACCTCAATGAGTTTCCCATCAGGCAGCCATTTCAGCAAGTGTGTTG 676

DB 972 TCACTGGCTTCTTGAACCTCAATGAGTTTCCCATCAGGCAGCCATTTCAGCAAGTGTGTTG 1031

QY 677 TCAGTAAATCTGAATATCCGTTGGGACAGGTACCCACCATTCGTCTCGGCACCGGCTACG 736

DB 1032 TCAGTAAATCTGAATATCCGTTGGGACAGGTACCCACCATTCGTCTCGGCACCGGCTACG 1091

DB 732 CCACAGTGCAGCGGAACGAAAGAGGTGGAGCAGGAGAGCCAGGACACCGGGGACATTA 791

QY 437 TCCAGAAAGGATTTCTAGACGTCTATTACATCTGACCTCGAACACCATGATGGCATAG 496

DB 792 TCCAGAAAGGATTTCTAGACGTCTATTACATCTGACCTCGAACACCATGATGGCATAG 851

QY 497 AATGGGTCCATCGCTTTTGTCTCAGCGGGGCTTTGTGATGAAACACAGCTCAGACATGT 556

DB 852 AATGGGTCCATCGCTTTTGTCTCAGCGGGGCTTTGTGATGAAACACAGCTCAGACATGT 911

QY 557 TCATCAATGTTGACTATCTGACTCACTGCTTCTGAGAAAAACAGAACACAGGTTTTT 616

DB 912 TCATCAATGTTGACTATCTGACTCACTGCTTCTGAGAAAAACAGAACACAGGTTTTT 971

QY 617 TCACTGGCTTCTTGAACCTCAATGAGTTTCCCATCAGGCAGCCATTTCAGCAAGTGTGTTG 676

DB 972 TCACTGGCTTCTTGAACCTCAATGAGTTTCCCATCAGGCAGCCATTTCAGCAAGTGTGTTG 1031

QY 677 TCAGTAAATCTGAATATCCGTTGGGACAGGTACCCACCATTCGTCTCGGCACCGGCTACG 736

DB 1032 TCAGTAAATCTGAATATCCGTTGGGACAGGTACCCACCATTCGTCTCGGCACCGGCTACG 1091

QY 737 TGTCTTCTGGCGAGCTGGCGAGTCAGGTGTACATGTCTCCAAGAGCGTCCCATACATTA 796

DB 1092 TGTCTTCTGGCGAGCTGGCGAGTCAGGTGTACATGTCTCCAAGAGCGTCCCATACATTA 1151

QY 797 AACTGGAAGACGTGTTTGTGGGGCTCTGCCCTCGAAAGGCTCAACATCAGATTGGAGGAGC 856

DB 1152 AACTGGAAGACGTGTTTGTGGGGCTCTGCCCTCGAAAGGCTCAACATCAGATTGGAGGAGC 1211

QY 857 TCCACTCCAGCCGACCTTTTTTCCAGGGGCTTACCGCTTCTCCGTATGCTCTTTCAGGA 916

DB 1212 TCCACTCCAGCCGACCTTTTTTCCAGGGGCTTACCGCTTCTCCGTATGCTCTTTCAGGA 1271

QY 917 GGATCGTGGCTGCCACTTCATCAAGCTCGGACTCTCTTGGACTACTGCGAGGCTCTAG 976

DB 1272 GGATCGTGGCTGCCACTTCATCAAGCTCGGACTCTCTTGGACTACTGCGAGGCTCTAG 1331

QY 977 AGAATCCCGGGGGGAAGATTGTCGCGCTGCTGA 1011

DB 1332 AGAATCCCGGGGGGAAGATTGTCGCGCTGCTGA 1366

RESULT 14

E38419

LOCUS E38419 2775 bp DNA linear PAT 31-JAN-2002

DEFINITION Novel polypeptide.

ACCESSION E38419

VERSION E38419.1 GI:18626993

KEYWORDS JP 2000245464-A/1.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 2775)

AUTHORS Narimatsu, H., Ishiki, S., Togayachi, A. and Sasaki, K.

TITLE Novel polypeptide

JOURNAL Patent: JP 2000245464-A 1 12-SEP-2000;

OS KYOWA HAKKO KOGYO CO LTD

PN JP 2000245464-A/1

PD 12-SEP-2000

PF 25-FEB-1999 JP 1999047571

PR HISASHI NARIMATSU, SOICHIRO ISSHIKI, AKIRA TOGAYACHI, PI

KATSUTOSHI SASAKI

PC C12N15/09.A01K67/027.C12N1/21.C12N5/10.C12N9/10.C12P19/00.PC

C12P21/02.

PC C12P21/08.C12Q1/68.G01N33/53//C12N1/21.C12R1:185).(C12N5/10,

PC C12R1:91).

PC C12P21/02.C12R1:185).(C12P21/02.C12R1:91).C12N15/00.C12N5/00,

CC (C12N5/00.C12R1:91)

FEATURES		source	Location/Qualifiers		Location/Qualifiers	FH	Key
					(402)..(1331).	FT	CDS
		1..2775					
		/organism="Homo sapiens"					
		/mol_type="genomic DNA"					
		/db_xref="taxon:9606"					
ORIGIN							
Query Match		91.8%;	Score 928.6;	DB 6;	Length 2775;		
Best Local Similarity		99.6%;	Pred. No. 7e-247;				
Matches 931;		Conservative 0;	Mismatches 4;	Indels 0;	Gaps 0;		
77	QY	AGATGGCTTCCCGAAGATGAGTTGATGTATATCTGCTTCTGCTTCTGGTCTCTGGGGGCTCTTT	136				
400	DB	AAATGGCTTCCCGAAGATGAGTTGATGTATATTTGCTTCTGCTTCTGGTCTCTGGGGGCTCTTT	459				
137	QY	GTTTGTATTTAGCATGTACAGTCTAAATCTCTTCAAGAAACAGTCTCTTTGTTTACAAGA	196				
460	DB	GTTTGTATTTAGCATGTACAGTCTAAATCTCTTCAAGAAACAGTCTCTTTGTTTACAAGA	519				
197	QY	AAGACGGAACTTCTTAAGCTCCAGATACAGACTGCAGGAGACACCTCCCTTCTCTCG	256				
520	DB	AAGACGGAACTTCTTAAGCTCCAGATACAGACTGCAGGAGACACCTCCCTTCTCTCG	579				
257	QY	TCCTGCTGCTGACCTCATCCCAACACAGTTGGCTGAGGCGCATGCCATCCGGCAGACGT	316				
580	DB	TCCTGCTGCTGACCTCATCCCAACACAGTTGGCTGAGGCGCATGCCATCCGGCAGACGT	639				
317	QY	GGGGAAAGAGAGGACGGTGAAGGGAAGACAGCTGAAGACATTTCTCTCTCGGGGACCA	376				
640	DB	GGGGAAAGAGAGGATGTTGAAGGGAAGACAGCTGAAGACATTTCTCTCTCGGGGACCA	699				
377	QY	CCAGCAGTGCAGCGGAAACAAAGAGGTGGACCCAGAGAGCCAGCGACACGGGGACATTA	436				
700	DB	CCAGCAGTGCAGCGGAAACAAAGAGGTGGACCCAGAGAGCCAGCGACACGGGGACATTA	759				
437	QY	TCCAGAAGGATTTCTTAGACGCTATTACAATCTGACCTCGAACACCATGATGGGCATAG	496				
760	DB	TCCAGAAGGATTTCTTAGACGCTATTACAATCTGACCTCGAACACCATGATGGGCATAG	819				
497	QY	AATGGGTCATCGCTTTTGTCTCAGCGGGCGTTTGTGATGAAAAACAGACTCAGACATGT	556				
820	DB	AATGGGTCATCGCTTTTGTCTCAGCGGGCGTTTGTGATGAAAAACAGACTCAGACATGT	879				
557	QY	TCATCAATGTTGACTATCTGACTGAATCTCTGAAAGAAAAACAGAAACACAGGTTTT	616				
880	DB	TCATCAATGTTGACTATCTGACTGAATCTCTGAAAGAAAAACAGAAACACAGGTTTT	939				
617	QY	TCACCTGGCTTCTTGAACCTCAATGATTTCCCATCAGGCAGCCATTACAGCAAGTGGTTTG	676				
940	DB	TCACCTGGCTTCTTGAACCTCAATGATTTCCCATCAGGCAGCCATTACAGCAAGTGGTTTG	999				
677	QY	TCAGTAATCTGAATATCCGTGGGACAGGTACCCACCACTTCTGCTCCGGCACCCGCTACG	736				
1000	DB	TCAGTAATCTGAATATCCGTGGGACAGGTACCCACCACTTCTGCTCCGGCACCCGCTACG	1059				
737	QY	TGTTTTCTGGGCGAGCTGGCGAGTCAATGTCTCCAAGAGCGTCCCATACATTA	796				
1060	DB	TGTTTTCTGGGCGAGCTGGCGAGTCAATGTCTCCAAGAGCGTCCCATACATTA	1119				
797	QY	AACGGAGAGCGTGTCTGGGGCTCTGCCTCGAAAGGCTGAACATCAGATTGGAGGAGC	856				
1120	DB	AACGGAGAGCGTGTCTGGGGCTCTGCCTCGAAAGGCTGAACATCAGATTGGAGGAGC	1179				
857	QY	TCCACTCCAGCGACCTTTTTTTCAGGGGGCTTACGCTTCTCCGTATGCTCTTTCAGGA	916				
1180	DB	TCCACTCCAGCGACCTTTTTTTCAGGGGGCTTACGCTTCTCCGTATGCTCTTTCAGGA	1239				
917	QY	GGATCGTGGCGCTGCCACTTCATCAAGCCTCGGACTCTTTGGACTACTGCGAGGCTCTAG	976				
1240	DB	GGATCGTGGCGCTGCCACTTCATCAAGCCTCGGACTCTTTGGACTACTGCGAGGCTCTAG	1299				

Qy	977	AGAATTCGGGGGAAGATTGTCCGCTGTCTGA	1011
Db	1300	AGAATTCGGGGGAAGATTGTCCGCTGTCTGA	1334
RESULT 15			
LOCUS AV372061			
DEFINITION Homo sapiens beta-1,3-galactosyltransferase 5 (B3GALT5) mRNA,			
complete cds.			
ACCESSION AV372061			
VERSION AV372061.1			
KEYWORDS GI:38045970			
SOURCE Homo sapiens (human)			
ORGANISM Homo sapiens			
REFERENCE 1 (bases 1 to 2921)			
AUTHORS Dunn, C.A., Medstrand, P. and Mager, D.L.			
TITLE An endogenous retroviral long terminal repeat is the dominant promoter for human [beta]1,3-galactosyltransferase 5 in the colon			
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (22), 12841-12846 (2003)			
PUBMED 14534330			
REFERENCE 2 (bases 1 to 2921)			
AUTHORS Dunn, C.A., Medstrand, P. and Mager, D.L.			
TITLE Direct Submission			
JOURNAL Submitted (20-AUG-2003) Terry Fox Laboratory, BC Cancer Agency, 601 West 10th Avenue, Vancouver, BC V5Z 1L3, Canada			
FEATURES			
source			
1. .2921			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/chromosome="21"			
/map="21q22.2"			
1. .2921			
/gene="B3GALT5"			
593..1525			
/note="Beta3Gal-T5"			
/gene="B3GALT5"			
/codon_start=1			
/product="beta-1,3-galactosyltransferase 5"			
/protein_id="AA08910.1"			
/db_xref="GI:38045971"			
/translation="MAFPKRLMYICLLVLGALCLYFSMYSINLPFKQS FVYKKDGNF			
LKUPDTRQTPPELVLLVTSSSHKQLAERMAIRQTGKERNVKGKQLKTFLLGTSS			
AAETKEVDQSRHGDI IQKDFLDVYNYLTKTMGIEWHRFCQAA FVNKLTSDMF			
INVDYELLLKKNRTRFTFGFLKNEFPPIRQPFKFWKSVKSYPMWDRYPFGSGTG			
YVPSGDVASOVYVSKSPVYIKLEDVFVGLCLERLNIRLBELHSQPTFFPGGLRFSVC			
LFRIVACHFIKPTLLDYWQLENSRGDCPPV"			
ORIGIN			
Query Match			
Best Local Similarity 91.8%; Score 928.6; DB 9; Length 2921;			
Matches 931; Conservative 0; Mismatches 4; Indels 0; Gaps 0;			
Qy	77	AGATGGCTTCCCGAAGATGAGTTGATGTATATCTGCTTCTGCTTCTGGGGGCTCTTT	136
Db	591	AAATGGCTTCCCGAAGATGAGTTGATGTATATTTGCTTCTGCTTCTGGGGGCTCTTT	650
Qy	137	GTTTGTATTTAGCATGTACAGTCTAAATCTCTTCAAGAAACAGTCTCTTTTACAAGA	196
Db	651	GTTTGTATTTAGCATGTACAGTCTAAATCTCTTCAAGAAACAGTCTCTTTTACAAGA	710
Qy	197	AAGACGGAACTTCTTAAGCTCCAGATACAGACTGCAGGAGACACCTCCCTTCCTCG	256
Db	711	AAGACGGAACTTCTTAAGCTCCAGATACAGACTGCAGGAGACACCTCCCTTCCTCG	770
Qy	257	TCCTGCTGTTGACCTCATCCCAAAACAGTTGGCTGAGCGCATGCCATCCGGCAGAGCT	316
Db	771	TCCTGCTGTTGACCTCATCCCAAAACAGTTGGCTGAGCGCATGCCATCCGGCAGAGCT	830

Qy	317	GGGGAAAGAGAGAGCGTGAAGGAAACAGCTGAAGACATTTCTTCTCTCTGGGGACCA	376
Db	831	GGGGAAAGAGAGAGATGCTGAAGGAAAGCAGCTGAGACATTTCTTCTCTGGGGACCA	890
Qy	377	CCAGCAGTGCAGCGGAAACAAAGAGGTGGACCAAGGAGCGACACCGGGACATTTA	436
Db	891	CCAGCAGTGCAGCGGAAACAAAGAGGTGGACCAAGGAGCGACACCGGGACATTTA	950
Qy	437	TCCAGAGGATTTCTTAGACGCTCTATTACAATCTGACCCCTGAAGACCAATGATGGGCATAG	496
Db	951	TCCAGAGGATTTCTTAGACGCTCTATTACAATCTGACCCCTGAAGACCAATGATGGGCATAG	1010
Qy	497	AATGGGTCATCGCTTTTCTCAGCGGGCTTTGTGATGAAGAACAGACTCAGACATGT	556
Db	1011	AATGGGTCATCGCTTTTCTCAGCGGGCTTTGTGATGAAGAACAGACTCAGACATGT	1070
Qy	557	TCATCAATGTTGACTATCTGACTGAATCTTCTGAAGAAACAGACAAACCAAGGTTTT	616
Db	1071	TCATCAATGTTGACTATCTGACTGAATCTTCTGAAGAAACAGACAAACCAAGGTTTT	1130
Qy	617	TCATCGGCTTCTTGAAACTCAATGAGTTTCCCATCAGGCAGCCATTCAGCAAGTGGTTTG	676
Db	1131	TCATCGGCTTCTTGAAACTCAATGAGTTTCCCATCAGGCAGCCATTCAGCAAGTGGTTTG	1190
Qy	677	TCAGTAAATCTGAATATCCGTGGGACAGGTACCCACCATTTCTGCTCCGGCACCGGCTACG	736
Db	1191	TCAGTAAATCTGAATATCCGTGGGACAGGTACCCACCATTTCTGCTCCGGCACCGGCTACG	1250
Qy	737	TGTTTCTCGGCGAGTGGCGAGTCAGGTGTACAATGTCTCCAAGAGCGTCCCATACATTTA	796
Db	1251	TGTTTCTCGGCGAGTGGCGAGTCAGGTGTACAATGTCTCCAAGAGCGTCCCATACATTTA	1310
Qy	797	AACTGGAAGACGTGTTGTGGGGCTCTGCCTCGAAAGCTGAACATCAGATTGGAGGAGC	856
Db	1311	AACTGGAAGACGTGTTGTGGGGCTCTGCCTCGAAAGCTGAACATCAGATTGGAGGAGC	1370
Qy	857	TCCACTCCAGCCGACCTTTTTCAGGGGGCTTACGCTTCTCGTATGCTCTTTCAGGA	916
Db	1371	TCCACTCCAGCCGACCTTTTTCAGGGGGCTTACGCTTCTCGTATGCTCTTTCAGGA	1430
Qy	917	GGATCGTGGCTTGGCACTTTCATCAAGCCTCGGACTCTCTTGGACTACTGGCAGGCTCTAG	976
Db	1431	GGATCGTGGCTTGGCACTTTCATCAAGCCTCGGACTCTCTTGGACTACTGGCAGGCTCTAG	1490
Qy	977	AGAATTCCCGGGGGAAGATTGTCCGCCCTGTCTGA	1011
Db	1491	AGAATTCCCGGGGGAAGATTGTCCGCCCTGTCTGA	1525

Search completed: September 22, 2005, 14:50:00
Job time : 4630.4 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2005, 12:09:12 ; Search time 4457.86 Seconds
(without alignments)
8632.613 Million cell updates/sec

Title: US-10-777-828-8
Perfect score: 1011
Sequence: 1 ccactcagcctctagcat.....aagatttcgcctgtctga 1011

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gsa1:*
9: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	931.4	92.1	933	9 AY419460	AY419460 Homo sapi
2	757.2	74.9	780	9 AY419461	AY419461 Pan trogl
3	525.6	52.0	5069	3 BC047347	BC047347 Mus muscu
4	517.6	51.2	927	9 AY419462	AY419462 Mus muscu
5	457.4	45.2	766	9 CB132434	CB132434 tigr-gss-
6	369.6	36.6	434	2 BF748181	BF748181 RC3-BN041
7	327.6	32.4	338	1 AJ003597	AJ003597 RC3-BN041
8	284.8	28.2	492	5 BP761359	BP761359 BP761359
9	260.6	25.8	447	6 CB546569	CB546569 AMGNUC:N
10	234	23.1	400	8 AZ768940	AZ768940 IM0569K09
11	185.6	18.4	1101	9 CNS05BT1	AL330238 Tetraodon
12	184.2	18.2	629	7 CR564113	CR564113 CR564113
13	182	18.0	981	9 AY401366	AY401366 Mus muscu
14	173	17.1	493	1 AJ507051	AJ507051 AJ507051
15	170	16.8	2705	3 AK035215	AK035215 Mus muscu
16	164	16.2	636	4 BG965086	BG965086 602829157
17	163.4	16.2	981	9 AY401364	AY401364 Homo sapi
18	161.4	16.0	981	9 AY401365	AY401365 Pan trogl
19	156.6	15.5	626	5 BX302370	BX302370 BX302370
20	154.8	15.3	901	4 BJ740020	BJ740020 BJ740020
21	144.8	14.3	1903	3 AK088407	AK088407 Mus muscu
22	144.2	14.3	632	9 CC554157	CC554157 CH240_460
23	143.2	14.2	996	9 AY402383	AY402383 Mus muscu
24	141	13.9	546	7 CR628501	CR628501 DKF2p459G

25	139.2	13.8	685	5 BUI33295	BUI33295 603122104
26	138.6	13.7	994	7 CF585163	CF585163 AGENCOURT
27	137	13.6	1514	3 AK003837	AK003837 Mus muscu
28	136	13.5	810	7 CF749272	CF749272 UI-M-HJO-
29	133.6	13.2	996	9 AY402381	AY402381 Homo sapi
30	132.4	13.1	682	6 BY704097	BY704097 BY704097
31	132.2	13.1	596	4 BJ501651	BJ501651 BJ501651
32	132	13.1	3564	3 BC028571	BC028571 Homo sapi
33	127	12.6	233	2 BB595953	BB595953 BB595953
34	123.6	12.2	811	7 CO396097	CO396097 AGENCOURT
35	123.4	12.2	444	4 BM403751	BM403751 zam6155_Z
36	123.2	12.2	638	4 BJ616136	BJ616136 BJ616136
37	122.6	12.1	879	5 BQ895238	BQ895238 AGENCOURT
38	122	12.1	842	5 BU350312	BU350312 603527845
39	121.2	12.0	792	7 CK030647	CK030647 AGENCOURT
40	121.2	12.0	1238	9 AY420414	AY420414 Homo sapi
41	120.6	11.9	652	5 BU627236	BU627236 UI-H-FGO-
42	119.4	11.8	697	6 CB233493	CB233493 AGENCOURT
43	118.6	11.7	1021	9 CNS044T4	AL274513 Tetraodon
44	118	11.7	1238	9 AY420416	AY420416 Mus muscu
45	118	11.7	2681	3 AK036141	AK036141 Mus muscu

ALIGNMENTS

RESULT 1
AY419460 933 bp DNA linear GSS 17-DEC-2003
LOCUS Homo sapiens B3GALT5 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY419460
VERSION AY419460.1 GI:39775417
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 933)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 933)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT These sequences were made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source location/Qualifiers
1..933
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>933
/gene="B3GALT5"
/locus_tag="HCM6898"

Query Match 92.1%; Score 931.4; DB 9; Length 933;
Best Local Similarity 99.9%; Pred. No. 1.2e-265;
Matches 932; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 79 ATGCGTTTCCGAGATGAGATTCATATATCGCTTCGTGGGGCTTTGT 138
|||||

Db 1 ATGGCTTTCCCGAAGATGAGATTGATGTATATTTTGCCTCTCTGGTCTCTGGGGCTCTTTGT 60
Qy 139 TTGTATTTTACAGCTACAGCTAAATCCTTTCAAGAACAGTCTTTTGTTTTCAAGAAA 198
Db 61 TTGTATTTTACAGCTACAGCTAAATCCTTTCAAGAACAGTCTTTTGTTTTCAAGAAA 120
Qy 199 GACGGAACTCTCTTAAGCTCCAGATACAGACTGCAGGACAGACCTCCCTCTCTCGTC 258
Db 121 GACGGAACTCTCTTAAGCTCCAGATACAGACTGCAGGACAGACCTCCCTCTCTCGTC 180
Qy 259 CTGTCTGTGACTCATCCCAAAACAGTTGGCTGAGCGCATGGCCATCCGGCAGACGTGG 318
Db 181 CTGTCTGTGACTCATCCCAAAACAGTTGGCTGAGCGCATGGCCATCCGGCAGACGTGG 240
Qy 319 GGGAAAGAGAGGCGGTGAAGGAAAGAGCTGAAGACATTTCTCTCTCTGGGACCAACC 378
Db 241 GGGAAAGAGAGGCGGTGAAGGAAAGAGCTGAAGACATTTCTCTCTCTGGGACCAACC 300
Qy 379 AGCAGTCAGCGGAAACAAAGAGAGTGACAGGAGAGCCAGCGACACGGGACATTTATC 438
Db 301 AGCAGTCAGCGGAAACAAAGAGAGTGACAGGAGAGCCAGCGACACGGGACATTTATC 360
Qy 439 CAGAAGGATTTCTTAGAGCTCTATTACATCTGACCTCTGAAGACCATGATGGGCATAGAA 498
Db 361 CAGAAGGATTTCTTAGAGCTCTATTACATCTGACCTCTGAAGACCATGATGGGCATAGAA 420
Qy 499 TGGGTCCATCGCTTTTGTCTCAGCGCGCTTTGTGATGAAACAGACTCAGACATGTTTC 558
Db 421 TGGGTCCATCGCTTTTGTCTCAGCGCGCTTTGTGATGAAACAGACTCAGACATGTTTC 480
Qy 559 ATCAATGTTGACTATCTGACTGAACTCTCTGAAGAAACAGACCAACAGGTTTTC 618
Db 481 ATCAATGTTGACTATCTGACTGAACTCTCTGAAGAAACAGACCAACAGGTTTTC 600
Qy 619 ACTGGCTCTTTGAAACTCAATGAGTTTCCCATCAGGACGACCATTCAGCAAGTGGTTGTC 678
Db 541 ACTGGCTCTTTGAAACTCAATGAGTTTCCCATCAGGACGACCATTCAGCAAGTGGTTGTC 600
Qy 679 AGTAAATCTGAATATCCGTGGGACAGTACCCACCATTTCTGCTCGGACCGGCTACGTG 738
Db 601 AGTAAATCTGAATATCCGTGGGACAGTACCCACCATTTCTGCTCGGACCGGCTACGTG 660
Qy 739 TTTTCTGGCGACGTGGCGAGTCAGGTGACAAATGTTCTCAAGAGCGTCCCATACATTTAA 798
Db 661 TTTTCTGGCGACGTGGCGAGTCAGGTGACAAATGTTCTCAAGAGCGTCCCATACATTTAA 720
Qy 799 CTGGAAGACGTGTTTGTGGGCTCTGCCCTCGAAAGGCTGAACATCAGATTGGAGGAGCTC 858
Db 721 CTGGAAGACGTGTTTGTGGGCTCTGCCCTCGAAAGGCTGAACATCAGATTGGAGGAGCTC 780
Qy 859 CACTCCAGCGACCTTTTTCAGGGGGCTTACGCTTCTCGGTATGCTCTTTCAGGAGG 918
Db 781 CACTCCAGCGACCTTTTTCAGGGGGCTTACGCTTCTCGGTATGCTCTTTCAGGAGG 840
Qy 919 ATCTGTGGCTCCCACTTCAATCAAGCTCGGACTCTCTTGGACTACTGCGGCTCTAGAG 978
Db 841 ATCTGTGGCTCCCACTTCAATCAAGCTCGGACTCTCTTGGACTACTGCGGCTCTAGAG 900
Qy 979 AATTCCCGGGGGGAAGATTGTCCGCTCTCTGA 1011
Db 901 AATTCCCGGGGGGAAGATTGTCCGCTCTCTGA 933

RESULT 2
LOCUS AY419461 780 bp DNA linear GSS 17-DEC-2003
DEFINITION Pan troglodytes B3GALT5 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY419461
VERSION AY419461.1 GI:39775418
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1 (bases 1 to 780)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarimal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 780)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarimal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
These sequences were made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..780
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>780
/gene="B3GALT5"
/locus_tag="HCM6898"
ORIGIN
Query Match 74.9%; Score 757.2; DB 9; Length 780;
Best Local Similarity 98.1%; Pred. No. 9.4e-214;
Matches 765; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
Qy 79 ATGGCTTTCCCGAAGATGAGATTGATGTATATTTTGCCTCTCTGGTCTCTGGGGCTCTTTGT 138
Db 1 ATGGCTTTCCCGAAGATGAGATTGATGTATATTTTGCCTCTCTGGTCTCTGGGGCTCTTTGT 60
Qy 139 TTGTATTTTACAGCTACAGCTAAATCCTTTCAAGAACAGTCTTTTGTTTTCAAGAAA 198
Db 61 GTGTATTTTACAGCTACAGCTAAATCCTTTCAAGAACAGTCTTTTGTTTTCAAGAAA 120
Qy 199 GACGGAACTCTCTTAAGCTCCAGATACAGACTGCAGGACAGACCTCCCTCTCTCGTC 258
Db 121 GACGGAACTCTCTTAAGCTCCAGATACAGACTGCAGGACAGACCTCCCTCTCTCGTC 180
Qy 259 CTGTCTGTGACTCATCCCAAAACAGTTGGCTGAGCGCATGGCCATCCGGCAGACGTGG 318
Db 181 CTGTCTGTGACTCATCCCAACAGTGGCTGAGCGCATGGCCATCCGGCAGACGTGG 240
Qy 319 GGGAAAGAGAGGACGTTGAAGGAAAGAGCTGAAGACATTTCTCTCTCTGGGACCAACC 378
Db 241 GGGAAAGAGAGGACGTTGAAGGAAAGAGCTGAAGACATTTCTCTCTCTGGGACCAACC 300
Qy 379 AGCAGTCAGCGGAAACAAAGAGAGTGACAGGAGCCAGCGACACGGGACATTTATC 438
Db 301 AGCAGTCAGCGGAAACAAAGAGAGTGACAGGAGCCAGCGACACGGGACATTTATC 360
Qy 439 CAGAAGGATTTCTTAGAGCTCTATTACATCTGACCTCTGAAGACCATGATGGGCATAGAA 498
Db 361 CAGAAGGATTTCTTAGAGCTCTATTACATCTGACCTCTGAAGACCATGATGGGCATAGAA 420
Qy 499 TGGGTCCATCGCTTTTGTCTCAGCGCGCTTTGTGATGAAACAGACTCAGACATGTTTC 558
Db 421 TGGGTCCATCGCTTTTGTCTCAGCGCGCTTTGTGATGAAACAGACTCAGACATGTTTC 480
Qy 559 ATCAATGTTGACTATCTGACTGAACTCTCTGAAGAAACAGACCAACAGGTTTTC 618
Db 481 ATCAATGTTGACTATCTGACTGAACTCTCTGAAGAAACAGACCAACAGGTTTTC 540
Qy 619 ACTGGCTCTTTGAAACTCAATGAGTTTCCCATCAGGACGACCATTCAGCAAGTGGTTGTC 678

Db 541 ACTGGCTCTTGAAGAACTCAATGAGTTTCCCATCAGCGAGCCCATTCAGTAAGTGGTTTCTC 600
 QY 679 AGTAATCTGTAATATCCGTGGGACAGGTACCCACCATTCCTGCTCGGACCGGCTACGTG 738
 Db 601 AGTAATCTGTAATATCCGTGGGACAGGTACCCGCCATTCCTGCTCGGACCGGCTACGTG 660
 QY 739 TTTTCTGGGAGCGTGGGAGTCAGGTGTACAAATGTCTCCAGAGCGTCCCATACATTTAAA 798
 Db 661 TTTTCCGGCGAGTGGCGAGTCAGGTGTACAAATGTCTCCAGAGCGTCCCATACATTTAAA 720
 QY 799 CTGGAAGACGCTGTTGTGGGGCTCTGCTCGAAAGGCTGAACATCAGATTGGAGGAGCTC 858
 Db 721 CTGGAAGACGCTGTTGTGGGGCTCTGCTCGAAAGGCTGAACATCAGATTGGAGGAGCTC 780

RESULT 3

BC047347
 LOCUS BC047347 5069 bp mRNA linear HTC 02-JAN-2004
 DEFINITION Mus musculus UDP-Gal:betaGlcNAc beta 1,3-galactosyltransferase,
 polypeptide 5, mRNA (cDNA clone IMAGE:5365404).

ACCESSION

BC047347

VERSION

BC047347.1 GI:28611171

KEYWORDS

HTC.

SOURCE

Mus musculus

ORGANISM

Mus musculus

REFERENCE

1 (bases 1 to 5069)

AUTHORS

Straubeberg, R.D., Collins, F.S., Wagner, L., Shermen, C.M., Schuler, G.D.,
 Klausner, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Woxley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whitling, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalusz, D.E.,
 Schnerch, A., Schein, J.E., Jones, J. and Marra, M.A.

TITLE

Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences

JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

PUBMED

12477932

REFERENCE

2 (bases 1 to 5069)

AUTHORS

Straubeberg, R.

TITLE

Direct Submission

JOURNAL

Submitted (28-FEB-2003) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REMARK

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs-remail.nih.gov

Tissue Procurement: The Cepko Laboratory

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: amg@bcm.tmc.edu

Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulaeeged, H.,
 Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
 A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 53 Row: j Column: 9
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 2212432
 This clone has the following problem: no 5' EST match.

FEATURES

source

1..5069
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:5365404"
 /tissue_type="Eye, retina, mouse strain C57BL/6"
 /clone_lib="NIH MGC 94"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"

ORIGIN

Query Match 52.0%; Score 525.6; DB 3; Length 5069;
 Best Local Similarity 73.3%; Pred. No. 2e-144;
 Matches 704; Conservative 0; Mismatches 244; Indels 12; Gaps 2;
 QY 55 ATCATTTGGATTTTCTTCTTTTCAGATGGCTTTCCGAGAGATGAGATTGATGTATATCTGC 114
 Db 568 AACGTGAGCTCTTACTCGGCAAGATGGCTCACATGAAGACAAGCTAGTTTACGCCTCC 627
 QY 115 CTTCCTGGTTCTGGGGGCTCTTTTGTGTATTTTATAGCATGTACAGTCTAAATCCTTTTCAA 174
 Db 628 ATTCTCATGATGGGCGCACTCTGCTGTGTACTTCAGCATGG-----ATTCTTTTCA 678
 QY 175 GAACAGTCTTTTGTGTACAGAAG--ACGGGAATCTTCTTAAGCTCCAGATACAGAC 231
 Db 679 GAATCCCGTTTGTGTATTTTAAAGAAAGTACGCGGAAGTTCTTCAGATTCCGGATATAG 738
 QY 232 TGCAGGACAGACACCTCCCTCTCTGCTGTGTGAGCTCATCCCAAAACAGTTTGGCT 291
 Db 739 TGAAGCAGAAAGCGCTTCTGCTGTGTGAGCTCATCTCAGACAGAGCTGAGCTGAGC 798
 QY 292 GAGCGCATGGCCATCCCGGACAGACGCTGGGGGAAAGAGAGGACGGTGAAGGGAAGACAGCTG 351
 Db 799 GCTCGCATGGCCATCCCGAAGACGCTGGGGTAGAGAGACATCTGTGCAGGGGCCAACAGGTG 858
 QY 352 AAGACATTTCTTCTCGGGACCAACAGCATGTGAGCGGAAACAAAAGAGGTGAGACAG 411
 Db 859 AGGACCTTCTTCTTCTGGGACCTCCGACAGCACCAGGAGAGATGGACGCCAACCCCTG 918
 QY 412 GAGAGCCAGACAGCGGGACATTTATCAGAGAGATTTCTTAGAGCTTATTTACATCTG 471
 Db 919 GAGAGCGAGACAGCAGCGGACATTTATCCAGAGGACCTTCAAGGATGCTTACTTTCAACCTG 978
 QY 472 ACCCTGAAGACCATGATGGGATAGAAATGGGTCCATCGCTTTTGTCTCAGCGCGCGTTT 531
 Db 979 ACCCTGAAGACCATGATGGGTATGGAATGGGTCTACCACTTTTGTCTCAGACAGCTTAC 1038
 QY 532 GTGATGAAAAACAGACTCAGACATGTTCAATGTTGATCTGATCTGACGTGAACCTGTTCTG 591
 Db 1039 GTGATGAAAAACGACTCTGACATGTTTGTGAATGTTGGCTATCTGACCGGAACCTGCTGCTA 1098
 QY 592 AAGAAACAGAACCAACGAGTTTTCACCTGGCTTTTGAACCTCAATGAGTTTCCCATC 651
 Db 1099 AAGAAACAAACACGACAGGTTCTTACAGGCTACATAAAGCCGCCAGACTTTTCCCATC 1158
 QY 652 AGCGAGCATTCAGCAAGTGGTTTGTGATGAATATCTGAAATATCCGTGGGACAGGTACCCA 711
 Db 1159 CGGCAAGTTTCAACAGTGGTTTGTGATGAATATCCCTGGGACAGGTACCCA 1218
 QY 712 CCATTTCTGCTCCGGACCGGCTACGTTGTTTCTGGCGAGCTGGGAGTCAAGGTGTACAA 771
 Db 1219 CTTTTTGTCTCGGTACTGTTGTTATGTTCTTTTCCAGCGAGCTGGGATCCAAAGTATACAA 1278
 QY 772 GTCTCCAGAGCGTCCCATACATTTAAACTGGAAGACGTTGTTGTTGGGCTCTGCTCGAA 831
 Db 1279 GTCTCAGAGCGGTTCCGTTTCATCAAGCTGAGGATGTGTTTGTGGGCTCTGCTTGCC 1338

QY	832	AGGCTGAACATCAGATTGGAGGAGCTCCACTCCACGCGACCTTTTTTTCAGGGGGCTTA	891
Db	1339	AAAGCTAAAGATCGGCGGAGGAGCTGCACACAAACAGACCTTCTTCCCTGGCGGTTTA	1398
QY	892	CGCTTCTCCGTATGCCCTTTCAGAGAGATCGTGGCGCTGCCACTTTCATCAAGCCTCGGACT	951
Db	1399	CGCTTTTCCGTGTGCCGCTTTCAGAAATATGTGGCATGCCATTTATATGAAGCCCCAGGAC	1458
QY	952	CTCTTGGACTACTGGCAGGCTCTAGAGAAATTCCGGGGGGAAGATTGTCCGCTGTCTCTGA	1011
Db	1459	CTGCTCACTTACTGGCAGAGCACTGGAGACTGGAAGAACAGGACTGTCTCTGTCTGTGA	1518
RESULT 4			
LOCUS	AY419462	927 bp DNA linear	GSS 17-DEC-2003
DEFINITION	Mus musculus B3GALT5 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.		
ACCESSION	AY419462		
VERSION	AY419462.1	GI:39775419	
KEYWORDS	GSS.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1 (bases 1 to 927) Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.		
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios		
JOURNAL	Science 302 (5652), 1960-1963 (2003)		
PUBMED	14671302		
REFERENCE	2 (bases 1 to 927) Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.		
AUTHORS	Direct Submission		
TITLE	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA		
JOURNAL			
COMMENT	These sequences were made by sequencing genomic exons and ordering them based on alignment.		
FEATURES	Location/Qualifiers		
source	1..927 /organism="Mus musculus" /mol_type="genomic DNA" /db_xref="taxon:10090" <1..>927 /gene="B3GALT5" /locus_tag="HCM6898"		
gene			
ORIGIN			
Query Match	51.2%;	Score 517.6;	DB 9; Length 927;
Best Local Similarity	73.7%;	Pred. No. 2.5e-142;	
Matches	690; Conservative	0; Mismatches 234; Indels 12; Gaps 2;	
QY	79	ATGGCTTTCCGGAAGATGAGATTGATATATCTGCCTCTCTGTCTCGGGGCTCTTTGT	138
Db	1	ATGGCTCACATGAAGACAAAGGCTAGTTTACGCCCTCCATTCTCATGATGGCGCACCTGCG	60
QY	139	TTCTATTTTAGCATGTACAGTCTAAATCCTTTCAAGAACAGTCCCTTTGTTTACAGAAA	198
Db	61	TTGTACTTTCAGCATGG-----ATTCTTTCAGAGAACTCCCGTTGTTGTTTAAAGAAA	111
QY	199	G---AAGCGAACTTCTTTAAGCTCCACAGATACAGACTGCAGGCAGACACCTCCCTTCCTC	255
Db	112	AGTCACGGGNAAGTTCTCTCAGATTCCGATATACAGCTGCAAGCAGAGCGCTTTCCTG	171
QY	256	GTCCCTGCTGTGTACCTCATCCCAACAAACAGTTGGCTGAGCGCATGCGCAATCCGCGAGA	315
Db	172	GTCTGTCTGTGATCGTCATCTCAACAGCAGCTGGCGCTCGCATGGCCATCCGCAAGACG	231

QY	316	TGGGGGAAAAGAGAGGACGGTGAAGGGGAAAAGCAGCTGAAGACATTCTTCTCTCTCTGGGGACC	375
Db	232	TGGGGTAGAGAGACATCTGTGCAGGGCCAAACAGGTGAGGACCTTCTTCTTCTTGGGGACC	291
QY	376	ACCAGCAGTCACGCGGAAAACAAAGAGGTGGACACAGGAGGAGCCAGCGACGACGGGACACATT	435
Db	292	TCCGACAGCACCGGAGGAGATGGACGCCACAAACCTCGGAGGAGGAGCAGCACCGGACATT	351
QY	436	ATCCAGAAGGATTTCTTAGACGCTCTATTACAATCTGACCCCTGAAGAGCATTGATGGGCATA	495
Db	352	ATCCAGAAGGACTTCAAGGATGCTCTTCAACCTGACCCCTGAAAACCATGATGGGTATG	411
QY	496	GAATGGGTCATCGCTTTTGTCTCAGGGGGCGTTTGTGATGAAAAACAGACTCAGACATG	555
Db	412	GAATGGGTCCTACCACTTTTGTCTCTCAGACAGCTTACGCTGATGAAAAACGGAATCTTGACATG	471
QY	556	TTCATCAATCTGACATCTGACCTGACCTGCTTCTGAAGAAAAACAGAACCAACAGGTTT	615
Db	472	TTTGTGAATGTGGCTATCTGACGGAACTGCTGCTTAAAGAAAAACAAACGACCAAGGTTT	531
QY	616	TTCACCTGGCTTCTTGAACACTCAATGAGTTTCCCATCAGGCAGCCATTTCAGCAAGTGGTTT	675
Db	532	TTTCACAGGCTACATAAAGCCCCAGCACTTTCCTCCATCCGCGAAGAAGTTCAACAAGTGGTTT	591
QY	676	GTCAAGTAATCTGAATATCCGTGGGAGCAGGTACCCACCATCTTGTCTCGGACACCGGCTAC	735
Db	592	GTGAGTAAGTTCCGAATATCCCTGGGACAGGTACCCAGCTTTTGTCTGTGTAAGTGGTTAT	651
QY	736	GTGTTTTCTCGGACAGTGGCGAGTCAGGTGTACAAATGTCTCAAGAGCGTCCCATACATT	795
Db	652	GTCTTTTCCAGCGACGTGGCGATCCAAAGTATACAAAGTCTCAGAGAGGTTCCGTTCACT	711
QY	796	AAACTCGAAGACGCTGTTTGTGGGCTCTGCTTCGAAAGGCTGAACATCAGATTGGAGGAG	855
Db	712	AAGCTGGAGGATGTGTTTGTGGGCTCTGCTTGGCCAAAGCTAAAGATCCGCGCGGAGGAG	771
QY	856	CTCACTCCACGCGACCTTTTTCAGGGGGCTTACGCTTCTCGTATGCTCTTCAGG	915
Db	772	CTCGACACCAAAACAGACCTTCTTCCCTGGCGGTTTACGCTTTTCGTTGCGCTTTCAG	831
QY	916	AGGATCGTGCCCTGCCACTTTCATCAAGCCTCGGAGCTCTTTGGAGTACTGGCAGGCTCTA	975
Db	832	AAATTGTGSCATGCCATTTTATGAAGCCCAAGACCTGCTCACTTACTGGCAACGACTG	891
QY	976	GAGAAATCCCGGGGGGAAGATTGTCGCTGTCTGA	1011
Db	892	GAGAACTCGAAAGAACAGGAGCTGCCCTGCTCTGA	927
RESULT 5	CE132434/c		
LOCUS	tigr-gss-dog-17000326239815	766 bp	DNA linear
DEFINITION	genomic survey sequence.		GSS 25-SEP-2000
ACCESSION	CE132434		
VERSION	CE132434.1	GI:35233692	
KEYWORDS	GSS.		
SOURCE	Canis familiaris (dog)		
ORGANISM	Canis familiaris		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.		
TITLE	1. (bases 1 to 766)		
JOURNAL	Kirkness, E.F., Bafna, V., Halpern, A.L., Levy, S., Remington, K.,		
MEDLINE	Rusch, D.B., Deicher, A.L., Pop, M., Wang, W., Fraser, C.M. and		
PUBMED	Venter, J.C.		
COMMENT	The dog genome: survey sequencing and comparative analysis		
	Science 301 (5641), 1698-1903 (2003)		
	22875432		
	14512627		
	Contact: Kirkness EF		
	The Institute for Genomic Research		
	Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive		


```

QY 876 TTTTCCAGGGGCTTAACGCTTTCGGTATGCTCTTTCAGAGGATCGTGGCTGCCACTT 935
Db 194 TTTTCCCGGGGCTTACGCTTTCGGTATGCTCTTTCAGAGGATCGTGGCTGCCACTT 135
QY 936 CATCAAGCTCGGACTCTCTTTGGACTACTTGGCAGGCTCTAGAGAATTCCTGGGGGGGAAGA 995
Db 134 CATCAAGCTCGGACTCTCTTTGGACTACTTGGCAGGCTCTAGAGAATTCCTGGGGGGGAAGA 75
QY 996 TTGTCCGCTGTCTGA 1011
Db 74 TTGTCCGCTGTCTGA 59

RESULT 7
AJ003597/c AJ003597 338 bp mRNA linear EST 04-DEC-1997
LOCUS AJ003597 Selected chromosome 21 cDNA library Homo sapiens cDNA
DEFINITION clone MPIp19-12J9, mRNA sequence.
ACCESSION AJ003597
VERSION AJ003597.1 GI:2578270
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 338)
TITLE Szulzewsky, I., Hunt, E., Nguyen, M., Korn, B., Roehrdanz, B.,
JOURNAL Lehrach, H. and Yaspo, M.L.H.
COMMENT An integrated transcript map for the whole human chromosome 21
Contact: Yaspo, M.-L.,
Max Planck Institut fuer Molekulare Genetik
Innestrasse 73, D14195 Berlin-Dahlem, Germany.
FEATURES
Location/Qualifiers
source
1..338
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="21q"
/clone="MPIp19-12J9"
/clone_lib="Selected chromosome 21 cDNA library"

ORIGIN
Query Match 32.4%; Score 327.6; DB 1; Length 338;
Best Local Similarity 98.5%; Pred. No. 7e-86;
Matches 330; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 93 GATGAGATTGATGATATCTGCTTCTGCTTCTGGGGCTCTTTGTTTGTATTTTAGCAT 152
Db 335 GACGAGATTGATGATATTTGTTTCTGGTCTGGGGCTCTTTGTTTGTATTTTAGCAT 276
QY 153 GTACAGTCTAAATCTTTTCAAGAAGACAGTCTTTGTTTCAAGAAGACGGGAATCTTCT 212
Db 275 GTACAGTCTAAATCTTTTCAAGAAGACAGTCTTTGTTTCAAGAAGACGGGAATCTTCT 216
QY 213 TAAGCTCCAGATACAGACTGAGCAGACAGCTCCCTCTTCTGCTCTGCTGCTGACCTC 272
Db 215 TAAGCTCCAGATACAGACTGAGCAGACAGCTCCCTCTTCTGCTCTGCTGCTGACCTC 156
QY 273 ATCCCAAAACAGTTGGCTGAGCGCATCGGCATCGGCGAGACGTGGGGGAAGAGAGGAC 332
Db 155 ATCCCAAAACAGTTGGCTGAGCGCATCGGCATCGGCGAGACGTGGGGGAAGAGAGGAC 96
QY 333 GGTGAAGGGAAGCAGCTGAAGACATTTCTTCTCTGGGGGACCAACAGCAGTGCAGCGGA 392
Db 95 GGTGAAGGGAAGCAGCTGAAGACATTTCTTCTCTGGGGGACCAACAGCAGTGCAGCGGA 36
QY 393 AACAAAAGAGGTGGACCAAGGAGAGCCAGCAGCAGC 427
Db 35 AACAAAAGAGGTGGACCAAGGAGAGCCAGCAGCAGC 1

```

```

RESULT 8
BP761359 492 bp mRNA linear EST 10-JUL-2004
LOCUS BP761359 mouse (C57BL/6) pancreatic islet library with
DEFINITION recombination-based method Mus musculus cDNA clone mic03025 5',
mRNA sequence.
ACCESSION BP761359
VERSION BP761359.1 GI:50220057
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 492)
TITLE Nishimura, M., Yokoi, N., Miki, T., Horikawa, Y., Yoshioka, H.,
JOURNAL Takeda, J., Ohara, O. and Seino, S.
COMMENT Construction of a multi-functional cDNA library specific for mouse
pancreatic islets and its application to microarray
Unpublished (2004)
CONTACT: Susumu Seino
Division of Cellular and Molecular Medicine
Kobe University Graduate School of Medicine
7-5-1 Kusunoki-cho, Chuo-ku, Kobe, Hyogo 650-0017, Japan
Tel: 81-78-382-5360
Fax: 81-78-382-5370
Email: seino@med.kobe-u.ac.jp.
Location/Qualifiers
source
1..492
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="mic03025"
/sex="male"
/tissue type="pancreatic islet"
/dev stage="adult"
/clone_lib="mouse (C57BL/6) pancreatic islet library with
recombination-based method"

ORIGIN
Query Match 28.2%; Score 284.8; DB 5; Length 492;
Best Local Similarity 76.5%; Pred. No. 4.6e-73;
Matches 349; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

QY 201 CGGAACCTCTTAAGCTCCAGATACAGACTGAGCAGCAGACCTCCCTTCCTGCTCT 260
Db 24 CGGGAAGTCTCTTCAGATTCGGATATAGACTGCAAGCAGAAAGCCGCTTTCTCTGGTGT 83
QY 261 GCTGTGACCTCATCCCAACAAACAGTTGGCTGAGCGCATGGCCATCGGCGAGACGTGGGG 320
Db 84 GCTGTGACCTCATCTCAAGCAGCTGGCGCTCGCATGGCCATCGCAAGACGTGGGG 143
QY 321 GAAAGAGAGGACGGTGAAGGAAAGCAGCTGAAGACATTTCTTCTCTGGGGAACCAAC 380
Db 144 TAGAGAGACATCTCTGCGAGGSCCAACAGGTGAGGACCTTCTTCTCTGCGGACCTCGGA 203
QY 381 CAGTGCAGCGAAACAAAGAGGTGGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 440
Db 204 CAGCAGCAGGAGATGGAACGACCAACCTTGAGAGCGAGCAGCAGCAGCAGCAGCAGCAGC 263
QY 441 GAAGGATTTCTAGACGCTTATTACATCTGACCTGAAGACCATGATGGGCATAGAAATG 500
Db 264 GAAGGATTTCAAGGATGCTTACTTCACTGACCTGAAACCATGATGGGTATGGAAATG 323
QY 501 GGTCATCGCTTTTGTCTCAGGGGGGTTTGTGATGAAAAACAGCTCAGACATGTTTCAAT 560
Db 324 GGTCTACACATTTTGTCTCAGACAGCTTACGTGATGAAACCGGACTCTGACATGTTTGT 383
QY 561 CAATGTTGATCTGACTGAACCTGCTTCTGAAGAAAAACAGAACCAACAGGTTTTTTCAC 620
Db 384 GAATGTTGGCTTTCTGACCGAACTGCTGTTTGAAGAAAAACCAACAGCAGGTTCTTTCAC 443
QY 621 TGGCTTCTTGAAACTCAATGAGTTTCCCATCAGGCA 656

```



```
|||||
444 AGGCTACATTAAAGCCCCACGACTTTTCCATTCGGCA 479

RESULT 9
CB546569
LOCUS CB546569 447 bp mRNA linear EST 01-APR-2003
DEFINITION AMGNNUC:NRHYS-00333-H3-A W Rat hypothalamus (10471), Rattus
norvegicus cDNA clone nrhys-00333-h3 5', mRNA sequence.
ACCESSION CB546569
VERSION CB546569
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
REFERENCE 1 (bases 1 to 447)
AUTHORS Amgen EST Program.
TITLE Amgen Rat EST Program
JOURNAL Unpublished (2003)
COMMENT Contact: Dan Fitzpatrick
Amgen, Inc
One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 00333 row: h column: 3.
Location/Qualifiers
1. 447
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone_lib="nrhys-00333-h3"
/clone_lib="W Rat hypothalamus (10471)"
/notes="Vector: pSPOR1; Site 1: SalI; Site 2: NotI; W Rat
hypothalamus adult female Wistar rat avg. Insert size 2.3
kb fraction 6 and 7"

FEATURES
source
ORIGIN
Query Match 25.8%; Score 260.6; DB 6; Length 447;
Best Local Similarity 76.6%; Pred. No. 7.3e-66;
Matches 333; Conservative 0; Mismatches 99; Indels 3; Gaps 1;

QY 156 CAGTCTAAATCCTTTCAAGAACAGTCTCTTTGTTTACAGAAA---GACGGGAATCTCT 212
DB 13 CAGCATGAGTCTTTTCCAAAGAACTCCCTTTGTTCTTCAAGAAAGTCAAGGAGTCTCT 72
QY 213 TAAGTCTCCAGATACAGATGCGAGCAGACACCTCCTTCTCTGCTCTGCTGTTGACCTC 272
DB 73 TCAGCTCCGGAGATAGACTGCAAGCAGAACGCGCTTTCTCTGCTGCTGTTGACGCTC 132
QY 273 ATCCCAAAACAGTTGGCTGAGCGCATGCGCAGATCGCGCAGAGCTGGGGGAAAGAGAGGAC 332
DB 133 ATCCCAACAGCAGTGGCGCTGCGATGCGCATGCGCAGAGAGCTGGGGTAGAGACATC 192
QY 333 GGTGAAGGGAAGCAGCTGAAGACATTTCTTCTCTGGGGACCCACAGCAGTGCAGCGGA 392
DB 193 GGTGACAGGGCCAAACCCGTTGAGGACCTTCTTCTCTGGGGTCTCTCAGACAGCAGCGGA 252
QY 393 AACAAAAGGTGGACACAGGAGAGCAGCAGGACACGCGGACATTTATCCAGAGGATTTCT 452
DB 253 CATGACGCTACAGCCCTGGAGAGGAGCAGCAGCAGTGCATCATCCAGAGGATTTCAA 312
QY 453 AGACGCTCTATTACATCTGACCTCAAGACCATGATGGCATAGATGGTCCATCGCTT 512
DB 313 GGACGCTATTTCACATTTGACCTGAAGACCATGATGGGTATGGATGGGTCTACCACTT 372
QY 513 TTGTCCTCAGCGCGGTTTGTGATGAACACAGACTCAGACATGTTTCATCAATGTTGACTA 572
DB 373 TTGTCCTCAGACCGCTTATGATGATAAAGCAGACTCAGACATGTTTGTGATGTTGGCTA 432
QY 573 TCTGACTGAATGCT 587
DB 433 TCTGACCGAATGCT 447

RESULT 10
A2768940/c
LOCUS A2768940 400 bp DNA linear GSS 16-FEB-2001
DEFINITION IM0569K09F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0569K09 F, genomic survey sequence.
ACCESSION A2768940
VERSION A2768940.1 GI:12888559
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 400)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
Niederhausern,A. and Wright,D.,Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
JOURNAL
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0569 row: K column: 09
Seq primer: CGTTGTAACACGACGCCAGT
Class: plasmid ends
High quality sequence stop: 400.
Location/Qualifiers
1. 400
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_lib="UUGC1M0569K09"
/sex="Male"
/lab_hosts="E. Coli strain XL10-Gold, Tl-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/notes="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (gi|4732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid RI. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
```


Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@anger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS SEQUENCE ID: ThDA01n08.plksp6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Nigel Garrett.
Seq primer: SP6.

FEATURES
Location/Qualifiers
1..629
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="THDA01n08"
/dev stages="tailbud head (stage 28-30)"
/lab host="Escherichia coli DH10B."
/clone lib="XGC-tailbud-head"
/note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA was oligo dT primed from sug of poly A+ RNA from tailbud head. EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the 5' end and NotI at the 3' end."

ORIGIN
Query Match 18.2%; Score 184.2; DB 7; Length 629;
Best Local Similarity 65.1%; Pred. No. 4.8e-43;
Matches 287; Conservative 0; Mismatches 153; Indels 1; Gaps 1;

QY 529 TTTGTGATGAAAAACAGACTCAGACATG-TTCATCAATGTTGACTATCTGACTGAACTGCT 587
DB 1 TTTGTGATGAGACCGACTCAGACATGTTTGTGTAACACCTTCTACCTGGTCCAACTGCT 60
QY 588 TCTGAAGAAAAACAGAACACAGGTTTTTCACTGGCTTTCTGAAACTCAATGAGTTTCC 647
DB 61 GGCAAGAAAAAACAGTCTTCTAATTTTTTACTGGTTTCTCAAACCTGAACGAGTACCC 120
QY 648 CATAGGAGCCATTCACCAAGTGTGTCAGTAAATCTGAATATCCGTGGACAGTA 707
DB 121 GATAAGGAATATCTTCAGCAAGTGATGCCAGTAAAGGGAATACCCAGGCGGCAAGTA 180
QY 708 CCCACCATCTGCTCCGGCACCGGTAGTGTGTTCTGGCAGCTGGCGAGTCAGGTGA 767
DB 181 CCCTCCATTTGTTGGGAGACTGATAGTCTTTCTGTAGACGTGCGCAAAAGATCCA 240
QY 768 CAATGTCTCCAGAGCGTCCCATACATTAATTAAGTGAAGACGTGTTTGTGGGGCTCTGCT 827
DB 241 CAACATCTCCAGCAGTGGCGTTTTCAAACTGGAGACGTCTATTGTTGGGCTATGCT 300
QY 828 CGAAGGCTGAACATCAGATTGGAGAGCTCCACTCCCGCGGACCTTTTTCGAGGGGG 887
DB 301 TGACATATTGACATCTCACTTGGAGAACTTCATACAGAGCAGACATTTCTTTGCGAGAG 360
QY 888 CTTAGGCTTCTCCGTATGCTCTTCAGGAGGATGTTGGCTGCCACTTTCATCAGCCCTCG 947
DB 361 CGAGTCATCTCCGTTTGGAAATACAGTAAGTGTGACGTCCTCCCAAGTGAACCCATA 420
QY 948 GACTCTCTGGACTACTGGCA 968
DB 421 TGAGACATTTGATCTGAA 441

RESULT 13
AY401366
LOCUS
DEFINITION Mus musculus B3GALT1 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY401366
VERSION AY401366.1 GI:39757355
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 981)

AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 981)
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES Location/Qualifiers
source
1..981
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/gene="B3GALT1"
/locus_tag="HCM0866"

gene

ORIGIN
Query Match 18.0%; Score 182; DB 9; Length 981;
Best Local Similarity 58.2%; Pred. No. 2.6e-42;
Matches 361; Conservative 0; Mismatches 250; Indels 9; Gaps 2;

QY 230 ACTCAGCGACACACCTCCCTTCCTCGTCTGCTGGTGCACCTCATCCCAACAAAGTTGG 289
DB 215 AGTCGGAGAAAAATCTCTTTCTCGTGAATCTTAGCACACACACAAGGAATTG 274
QY 290 CTGAGCGCATGGCCATCCGGCAGACGTGGGGGAAAGAGAGGACGGTGAAGGAAAGCAGC 349
DB 275 ATGCCCGCCAGGCATCCGGGAGACATGGGGGATGAAAACAACCTTCAAGGGATCAAGA 334
QY 350 TGAAGACATTTCTCTCTGGGGACCAACAGCAGTGGCAGCGGNAACAAGAGTGGACC 409
DB 335 TAGCCACACTTTCTCTCTGGGCAAAATGCTGATCTCTGAACAGATGGTGGAGC 394
QY 410 AGGAGCCAGCGACAGCGGACATATCCAGAAAGGATTTCTAGACGCTATTACAATC 469
DB 395 AAGAGAGCCAGATCTTCATGACATCATCGTGGAGGACTTCATTGACTCTTACCACATC 454
QY 470 TGACCCCTGAAGACCATGATGGGCATAGAAATGGGTCCATCGCTTTTGTCTCAGGGCGGT 529
DB 455 TCACCCCTCAAAACCTTAATGGGATGAGATGGGTGCGACTTCTGTTCAAAAGGCCAAGT 514
QY 530 TTGTGATGAAACAGATCAGACATGTTCAATGTTGACTATCTGACTGACTGACTGCTTC 589
DB 515 ACGTGATGAAACCGACAGTGCACATTTTGTGAAACATGGAACAACCTTATTATAAATCC 574
QY 590 TGA-----GAAAAACAGAACCAACAGGTTTTTCACTGGCTTCTTGAACCTCAATGAGT 643
DB 575 TGAACCCCTTACCAAGCCAGAGAGAGATTTTCACTGGTGTAGCTCA---TCAAGCGG 631
QY 644 TTCCCATCAGCAGCCATTTCAGCAAGTGGTTTGTGAGTAAATCTGAATATCCGTGGGACA 703
DB 632 GGCCAACTCAGGATGTCGCGAGTAAGTGGTATATGCTTAGAGATTGTGACCTGACAGCA 691
QY 704 GGTACCAACCATTTCTGCTCCGGCACCGGCTACGTGTTTTCTGGCGACGTGGGAGTCAAG 763
DB 692 ACTACCCACCGTCTCTGTTTCAGGAGCTGGCTATATCTTTCCGCTGATGTGGCTGAAC 751
QY 764 TGTACAAATGCTCCCAAGAGCGTCCCATACATTAACACTGGAAGAGCTGTTTGTGGGGCT 823
DB 752 TTTACAGACCTCGCTCCACAGCGGTGCTTTCATCTTGAAGATGTGTACGTGGGACTGT 811
QY 824 GCCTCGAAGGCTGAACATC 843

Db	 812 GTCTTCGAAAGCTTGCCATC 831
RESULT 14 AJ507051	
LOCUS	AJ507051 493 bp mRNA linear EST 17-SEP-2002
DEFINITION	AJ507051 Mus musculus BALB/c (Kahlem P) Mus musculus cDNA clone
ACCESSION	AJ507051
VERSION	AJ507051.1 GI:23034136
KEYWORDS	EST.
SOURCE	Mus musculus (house mouse)
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	1 Gitton,Y., Dahmane,N., Baik,S., Ruiz i Altaba,A., Neidhardt,L., Scholze,M., Herrmann,B.G., Kahlem,P., Ben Kahla,A., Schrinner,S., Yildirimman,R., Herwig,R., Leirach,H. and Yaspo,M.L. A gene expression map of human chromosome 21 orthologs in the mouse Nature (2002) In press Contact: Kahlem P Vertebrat Genomics Max-Planck Institute for Molec. Genet. Innestrasse 73, D-14195 Berlin, GERMANY.
AUTHORS	1 .493 /organism="Mus musculus" /mol_type="mRNA" /strain="BALB/c" /db_xref="taxon:10090" /clone="182" /clone_lib="Mus musculus BALB/c (Kahlem P)"
TITLE	
JOURNAL	
COMMENT	
FEATURES	
source	
ORIGIN	
Query Match	17.1%; Score 173; DB 1; Length 493;
Best Local Similarity	73.9%; Pred. No. 9,6e-40;
Matches 232; Conservative	0; Mismatches 81; Indels 1; Gaps 1;
QY	671 GGTTGTCTGAATAACTGAATATCGTGGGCACAGTA-CGCCACATTCTGTCCTCGGCACC 729
Db	 672 GGTTGTCTGAATAACTGAATATCGTGGGCACAGTACCCCACCTTTTGTCTCTGCTACT 111
QY	730 GCGTAGCTGTTTTCTGGCAGCGTGGCAGTACAGTGTACATGCTCTCCAGAGCGTCCCCA 789
Db	 112 GGTATGTCTTTTTCAGCGACGTGGCATCAAGPATACAACGCTCTCAGAGAGGTTTCCG 171
QY	790 TACATTAAACTGGAGAAGCTGTTCTGGGGCTCTGCTCGAAGGCTGAACATCAGATTG 849
Db	 172 TTCATCAAGCTGAGGAGTGTTGTGTTGGCTCCGCTTGGCCAAGCTTAAGATCCGGCCG 231
QY	850 GAGGAGCTCCATCCCGACCGACCTTTTTCAGGGGCTTTACGCTTCTCCGTTATGCTTC 909
Db	 232 GAGGAGCTGCACACCAAACAGACCTTCTTCCCTGGCGGTTTACGCTTTCCGTGTCGCCG 291
QY	910 TTCAGGAGGATCGTGGCTGCCACTTCATCAAGCTCCGACTCTCTTGGACTACTGGCAG 969
Db	 292 TTTCANNAAAATTTGTGGCATGCCATTTTATGAAGCCCCAGCACCTGCTCACTTACTGGCAA 351
QY	970 GCTCTAGAGAAATTC 983
Db	 352 GCCTGGAGAACTC 365
RESULT 15 AK035215	
LOCUS	AK035215 2705 bp mRNA linear HTC 03-APR-2004
DEFINITION	Mus musculus adult male urinary bladder cDNA, RIKEN full-length enriched library, clone:1953002C17 product:UDP-Gal:betaGalCNAC beta
ACCESSION	AK035215
VERSION	AK035215.1 GI:26084471
KEYWORDS	
SOURCE	
ORGANISM	
REFERENCE	
AUTHORS	
TITLE	
JOURNAL	
COMMENT	
FEATURES	
source	
ORIGIN	
Query Match	17.1%; Score 173; DB 1; Length 493;
Best Local Similarity	73.9%; Pred. No. 9,6e-40;
Matches 232; Conservative	0; Mismatches 81; Indels 1; Gaps 1;
QY	671 GGTTGTCTGAATAACTGAATATCGTGGGCACAGTA-CGCCACATTCTGTCCTCGGCACC 729
Db	 672 GGTTGTCTGAATAACTGAATATCGTGGGCACAGTACCCCACCTTTTGTCTCTGCTACT 111
QY	730 GCGTAGCTGTTTTCTGGCAGCGTGGCAGTACAGTGTACATGCTCTCCAGAGCGTCCCCA 789
Db	 112 GGTATGTCTTTTTCAGCGACGTGGCATCAAGPATACAACGCTCTCAGAGAGGTTTCCG 171
QY	790 TACATTAAACTGGAGAAGCTGTTCTGGGGCTCTGCTCGAAGGCTGAACATCAGATTG 849
Db	 172 TTCATCAAGCTGAGGAGTGTTGTGTTGGCTCCGCTTGGCCAAGCTTAAGATCCGGCCG 231
QY	850 GAGGAGCTCCATCCCGACCGACCTTTTTCAGGGGCTTTACGCTTCTCCGTTATGCTTC 909
Db	 232 GAGGAGCTGCACACCAAACAGACCTTCTTCCCTGGCGGTTTACGCTTTCCGTGTCGCCG 291
QY	910 TTCAGGAGGATCGTGGCTGCCACTTCATCAAGCTCCGACTCTCTTGGACTACTGGCAG 969
Db	 292 TTTCANNAAAATTTGTGGCATGCCATTTTATGAAGCCCCAGCACCTGCTCACTTACTGGCAA 351
QY	970 GCTCTAGAGAAATTC 983
Db	 352 GCCTGGAGAACTC 365
HTC; CAP trapper.	
Mus musculus (house mouse)	
Mus musculus	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
1 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999) 99279253 10349636	
2 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000) 20499374 11042159	
3 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Wataniki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000) 20530913 11076861	
4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)	
5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)	
6 (bases 1 to 2705)	
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Horii,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru.A., Toya,T., Yasunishi,A., Muramatsu.M. and Hayashizaki,Y. Direct Submission Submitted (16-JUN-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216) cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL:http://genome.gsc.riken.jp/ URL:http://fantom.gsc.riken.jp/. Location/Qualifiers	
COMMENT	
FEATURES	

Search completed: September 22, 2005, 16:55:53
Job time : 4465.86 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2005, 14:50:12 ; Search time 2624.46 Seconds
(without alignments)
2575.655 Million cell updates/sec

Title: US-10-777-828-8

Perfect score: 1011
Sequence: 1 ccacttcagcctctagcat.....aagatttcgcgtgtctga 1011

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 7400704 seqs, 3343079526 residues

Total number of hits satisfying chosen parameters: 14801408

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:**

- 1: /cgm2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgm2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgm2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgm2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgm2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgm2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgm2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgm2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgm2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgm2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgm2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgm2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgm2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgm2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgm2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgm2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgm2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgm2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgm2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
- 20: /cgm2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
- 21: /cgm2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
- 22: /cgm2_6/ptodata/1/pubpna/US10J_NEW_PUB.seq.*
- 23: /cgm2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
- 24: /cgm2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 25: /cgm2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 26: /cgm2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1011	100.0	1011	19	US-10-777-828-8
2	640	63.3	681	13	Sequence 8, Appli
3	640	63.3	681	13	Sequence 21208, A
4	640	63.3	681	13	Sequence 21209, A
5	640	63.3	681	13	Sequence 21210, A
6	640	63.3	681	17	Sequence 21208, A
7	640	63.3	681	17	Sequence 21209, A
					Sequence 21210, A

8	515.6	51.0	516	13	US-10-027-632-106210	Sequence 106210,
9	515.6	51.0	516	17	US-10-027-632-106210	Sequence 106210,
10	514	50.8	516	13	US-10-027-632-106208	Sequence 106208,
11	514	50.8	516	13	US-10-027-632-106209	Sequence 106209,
12	514	50.8	516	17	US-10-027-632-106208	Sequence 106208,
13	514	50.8	516	17	US-10-027-632-106209	Sequence 106209,
14	405.2	40.1	596	16	US-10-029-386-13121	Sequence 13121, A
15	161.8	16.0	2168	21	US-10-887-553A-603	Sequence 603, App
16	136.8	13.5	1965	21	US-10-764-420-1500	Sequence 1500, Ap
17	133.6	13.2	1358	20	US-10-370-715B-481	Sequence 481, App
18	133.6	13.2	1358	21	US-10-651-237-77	Sequence 77, Appl
19	133.6	13.2	1358	21	US-10-782-413-77	Sequence 77, Appl
20	133.6	13.2	1773	9	US-03-739-451-9	Sequence 9, Appli
21	133.6	13.2	2095	9	US-09-989-722-208	Sequence 208, App
22	133.6	13.2	2095	9	US-09-989-723-208	Sequence 208, App
23	133.6	13.2	2095	9	US-09-989-279-208	Sequence 208, App
24	133.6	13.2	2095	9	US-09-989-727-208	Sequence 208, App
25	133.6	13.2	2095	9	US-09-989-731-208	Sequence 208, App
26	133.6	13.2	2095	9	US-09-989-732-208	Sequence 208, App
27	133.6	13.2	2095	9	US-09-991-073-208	Sequence 208, App
28	133.6	13.2	2095	9	US-09-990-442-208	Sequence 208, App
29	133.6	13.2	2095	9	US-09-991-163-208	Sequence 208, App
30	133.6	13.2	2095	9	US-09-993-604-208	Sequence 208, App
31	133.6	13.2	2095	9	US-09-990-456-208	Sequence 208, App
32	133.6	13.2	2095	9	US-09-989-721-208	Sequence 208, App
33	133.6	13.2	2095	9	US-09-992-598-208	Sequence 208, App
34	133.6	13.2	2095	9	US-09-989-293A-208	Sequence 208, App
35	133.6	13.2	2095	9	US-09-989-735-208	Sequence 208, App
36	133.6	13.2	2095	9	US-09-990-444-208	Sequence 208, App
37	133.6	13.2	2095	9	US-09-991-181-208	Sequence 208, App
38	133.6	13.2	2095	9	US-09-989-730-208	Sequence 208, App
39	133.6	13.2	2095	9	US-09-990-436-208	Sequence 208, App
40	133.6	13.2	2095	9	US-09-993-687-208	Sequence 208, App
41	133.6	13.2	2095	10	US-09-989-734-208	Sequence 208, App
42	133.6	13.2	2095	10	US-09-997-653-208	Sequence 208, App
43	133.6	13.2	2095	10	US-09-989-724-208	Sequence 208, App
44	133.6	13.2	2095	10	US-09-989-728-208	Sequence 208, App
45	133.6	13.2	2095	10	US-09-990-441-208	Sequence 208, App

ALIGNMENTS

RESULT 1

US-10-777-828-8
; Sequence 8, Application US/10777828
; Publication NO. US20040142425A1
; GENERAL INFORMATION:
; APPLICANT: Clausen, Henrik
; APPLICANT: Amado, Margarita
; TITLE OF INVENTION: UDP-GALACTOSE: BETA-N-ACETYL-GLUCOSAMINE BETA 1,3
; TITLE OF INVENTION: GALACTOSYLTRANSFERASES, BETAGAL-T5
; FILE REFERENCE: 7188-157
; CURRENT APPLICATION NUMBER: US/10/777,828
; PRIOR FILING DATE: 2004-02-12
; PRIOR APPLICATION NUMBER: US/09/831,630
; PRIOR FILING DATE: 2001-05-10
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 1011
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (79)..(1008)
us-10-777-828-8

Query Match 100.0%; Score 1011; DB 19; Length 1011;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1011; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 CCACCTCAGCCTCCTAGCATAAACTAGACACATCCTCTTCGTTTGTAGGCTTAATCATTT 60

```

1 CCACCTCAGCCTCCTAGCATAAACTAGACACATCCTCATGCTTTTGGGCTCATCATTT 60
61 GGATTTTGTCTCTTTCAGATGGCTTTCGGAAGATGAGATTGATATATCTGCTTCTG 120
61 GGATTTTGTCTCTTTCAGATGGCTTTCGGAAGATGAGATTGATATATCTGCTTCTG 120
121 GTTCTGGGGGCTCTTGTGTTGTTATTTTATGATGATGATGATGATGATGATGATG 180
121 GTTCTGGGGGCTCTTGTGTTGTTATTTTATGATGATGATGATGATGATGATGATG 180
181 TCCCTTTGTTTACAGAAAGAGCGGAACTTCTTAACTCCAGATGATGATGATGATG 240
181 TCCCTTTGTTTACAGAAAGAGCGGAACTTCTTAACTCCAGATGATGATGATGATG 240
241 ACACCTCCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
241 ACACCTCCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
301 GCCATCCGGCAGACGTGGGGGAAAGAGAGGACGGTGAAGGAAAGAGAGAGAGAG 360
301 GCCATCCGGCAGACGTGGGGGAAAGAGAGGACGGTGAAGGAAAGAGAGAGAGAG 360
361 TTCTCTCTCTGGGGACCAACAGCAGTGCAGCGGAAACAAAGAGAGGAGGAGAG 420
361 TTCTCTCTCTGGGGACCAACAGCAGTGCAGCGGAAACAAAGAGAGGAGGAGAG 420
421 CGACACGGGGACATTAATCCAGAGGATTTCTAGACGCTTATTAATCTGACCTG 480
421 CGACACGGGGACATTAATCCAGAGGATTTCTAGACGCTTATTAATCTGACCTG 480
481 ACCATGATGGGCATAGATGGGTGCATCGCTTTTCTCTGCGGGGCTTTTGTGAT 540
481 ACCATGATGGGCATAGATGGGTGCATCGCTTTTCTCTGCGGGGCTTTTGTGAT 540
541 ACAGACTCAGACATGTTTCAATGTTGACTATCTGACTGAACTGCTTCTGAAAG 600
541 ACAGACTCAGACATGTTTCAATGTTGACTATCTGACTGAACTGCTTCTGAAAG 600
601 AGAACAAACAGGTTTTTCACTGGCTTCTGAAAGTCTGAAAGTCTGAAAGTCT 660
601 AGAACAAACAGGTTTTTCACTGGCTTCTGAAAGTCTGAAAGTCTGAAAGTCT 660
661 TTCAGCAAGTGGTGTGCTGATTAATCTGAAATCTGAAATCTGAAATCTGAA 720
661 TTCAGCAAGTGGTGTGCTGATTAATCTGAAATCTGAAATCTGAAATCTGAA 720
721 TCCGGCACCGGCTACGTTGTTTCTGGGACGCTGGGAGTCAAGTGTCAATGTT 780
721 TCCGGCACCGGCTACGTTGTTTCTGGGACGCTGGGAGTCAAGTGTCAATGTT 780
781 AGCGTCCCATATTAATTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
781 AGCGTCCCATATTAATTAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
841 ATCAGATGGAGGAGCTCCATCCCGAGCGGACCTTTTTCAGGGGGCTTACGCT 900
841 ATCAGATGGAGGAGCTCCATCCCGAGCGGACCTTTTTCAGGGGGCTTACGCT 900
901 GTATGCTCTCTCAGGAGGATCGTGGCTGCGCACTTTCATCAAGCCCTCGGACT 960
901 GTATGCTCTCTCAGGAGGATCGTGGCTGCGCACTTTCATCAAGCCCTCGGACT 960
961 TACTGGCAGGCTCTAGAGAAATCCCGGGGGAGAGATTGTCGCCCTGTCTGA 1011
961 TACTGGCAGGCTCTAGAGAAATCCCGGGGGAGAGATTGTCGCCCTGTCTGA 1011

```

RESULT 2
 US-10-027-632-21208
 ; Sequence 21208, Application US/10027632
 ; Publication No. US20020198371A1
 ; GENERAL INFORMATION:

```

; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: Polymorphisms in the Human Genome
; CURRENT FILING DATE: 2002-04-30
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21208
; LENGTH: 681
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-21208

Query Match      63.3%; Score 640; DB 13; Length 681;
Best Local Similarity 99.1%; Pred. No. 1.2e-201;
Matches 640; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 358 TTCTTCTCTCTCTGGGGACCAACAGCAGTGCAGCGGAAACAAAGAGAGGAGGAGGAGGAGC 417
Db 6 TTCTTCTCTCTCTGGGGACCAACAGCAGTGCAGCGGAAACAAAGAGAGGAGGAGGAGC 65
Qy 418 CAGCGACACCGGGACATTTATCCAGAGGATTTCTAGACGCTTATTAATCTGACCTG 477
Db 66 CAGCGACACCGGGACATTTATCCAGAGGATTTCTAGACGCTTATTAATCTGACCTG 125
Qy 478 AAGACCATGATGGGCATAGAAATGGGTCCATCGCTTTTGTCTCAGGGCGGCTTTGTGATG 537
Db 126 AAGACCATGATGGGCATAGAAATGGGTCCATCGCTTTTGTCTCAGGGCGGCTTTGTGATG 185
Qy 538 AAAACAGACTCAGACATGTTTCAATGTTGACTATCTGACTGAACTGCTTCTGAAGAAA 597
Db 186 AAAACAGACTCAGACATGTTTCAATGTTGACTATCTGACTGAACTGCTTCTGAAGAAA 245
Qy 598 AACAGAACACCGAGTTTTTTCACCTGGCTTCTTGAAGAACTCAATGAGTTTCCCATCAGG 657
Db 246 AACAGAACACCGAGTTTTTTCACCTGGCTTCTTGAAGAACTCAATGAGTTTCCCATCAG 305
Qy 658 CCATTTCAGCAAGTGGTGTGTCAGTAAATCTGAATPATCCGTGGGACAGGTACCCACCA 717
Db 306 CCATTTCAGCAAGTGGTGTGTCAGTAAATCTGAATPATCCGTGGGACAGGTACCCACCA 365
Qy 718 TGCTCCGGCACCGGCTACGTTTCTGGGACGCTGGGAGTCAAGTGTACAAATGCTCTCC 777
Db 366 TGCTCCGGCACCGGCTACGTTTCTGGGACGCTGGGAGTCAAGTGTACAAATGCTCTCC 425
Qy 778 AAGAGCGTCCCATACATTTAAACTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 837
Db 426 AAGAGCGTCCCATACATTTAAACTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 485
Qy 838 AACATCAGATTGGAGGAGCTCCACTCCCGAGCGGACCTTTTTCAGGGGGCTTACGCTTC 897
Db 486 AACATCAGATTGGAGGAGCTCCACTCCCGAGCGGACCTTTTTCAGGGGGCTTACGCTTC 545
Qy 898 TCCGATGCTCTTTCAGGAGGATCGTGGCTGCGCACTTTCATCAAGCCCTCGGACTCTCT 957
Db 546 TCCGATGCTCTTTCAGGAGGATCGTGGCTGCGCACTTTCATCAAGCCCTCGGACTCTCT 605
Qy 958 GACTACTGGCAGGCTCTAGAGAAATTCGGGGGGGAGAGATTGTCGCCG 1003

```

Db 606 GACTACTGCGAGGCTCTAGAGAAATCCCGGGGGAAGATTGGCCGC 651

RESULT 3
US-10-027-632-21209
; Sequence 21209, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21209
; LENGTH: 681
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-21209

Query Match 63.3%; Score 640; DB 13; Length 681;
Best Local Similarity 99.1%; Pred. No. 1.2e-201;
Matches 640; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 358 TTCTTCTCTCTGGGACCAACAGAGTGCAGCGGAAACAAAGAGGTGGACAGGAGGC 417
Db 6 TTCTTCTCTCTGGGACCAACAGAGTGCAGCGGAAACAAAGAGGTGGACAGGAGGC 65
Qy 418 CAGGACACAGCGGGACATTTATCCAGAGGATTTCTAGAGCTTATTACAATCTGACCTG 477
Db 66 CAGGACACAGCGGGACATTTATCCAGAGGATTTCTAGAGCTTATTACAATCTGACCTG 125
Qy 478 AAGACCATGATGGGCATAGAAATGGGTCCATCGCTTTTGTCTCAGCGCGGCTTGTGATG 537
Db 126 AAGACCATGATGGGCATAGAAATGGGTCCATCGCTTTTGTCTCAGCGCGGCTTGTGATG 185
Qy 538 AAAACAGACTCAGACATGTTTCAATCAATGTTGAACTCAATGAGTTTCCCATCAGGAG 597
Db 186 AAAACAGACTCAGACATGTTTCAATCAATGTTGAACTCAATGAGTTTCCCATCAGGAG 245
Qy 598 AACAGAACACAGGTTTTTCACTGGCTTCTTGAACCTCAATGAGTTTCCCATCAGGAG 657
Db 246 AACAGAACACAGGTTTTTCACTGGCTTCTTGAACCTCAATGAGTTTCCCATCAGGAG 305
Qy 658 CCATTCAGCAAGTGGTTGTGAGTAAATCTGAATATCCGTGGGACAGGTACCCACCATTC 717
Db 306 CCATTCAGCAAGTGGTTGTGAGTAAATCTGAATATCCGTGGGACAGGTACCCACCATTC 365
Qy 718 TGCTCCGACACCGGCTACGTTTCTGGGACAGTGGGAGTTCAGGTGTAACAATGTTCTCC 777
Db 366 TGCTCCGACACCGGCTACGTTTCTGGGACAGTGGGAGTTCAGGTGTAACAATGTTCTCC 425
Qy 778 AAGAGCTGCCATACATTAATAAAGAGAGTGTGTTGTGGGCTCTGCTCGAAGGCTG 837
Db 426 AAGAGCTGCCATACATTAATAAAGAGAGTGTGTTGTGGGCTCTGCTCGAAGGCTG 485

Qy 838 AACATCAGATTGGAGGAGCTCCACTCCAGCGGACCTTTTTTCCAGGGGCTTACGCTTC 897
Db 486 AACATCAGATTGGAGGAGCTCCACTCCAGCGGACCTTTTTTCCAGGGGCTTACGCTTC 545
Qy 898 TCCGTATGCTCTTCCAGGAGGATCGTGGCCCTGACATTCATCAAGCCCTCGGACTCTCTTG 957
Db 546 TCCGTATGCTCTTCCAGGAGGATCGTGGCCCTGACATTCATCAAGCCCTCGGACTCTCTTG 605
Qy 958 GACTACTGCGAGGCTCTAGAGAAATTCCTCGGGGGAAGATTGTCCGC 1003
Db 606 GACTACTGCGAGGCTCTAGAGAAATTCCTCGGGGGAAGATTGTCCGC 651

RESULT 4
US-10-027-632-21210
; Sequence 21210, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21210
; LENGTH: 681
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-21210

Query Match 63.3%; Score 640; DB 13; Length 681;
Best Local Similarity 99.1%; Pred. No. 1.2e-201;
Matches 640; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 358 TTCTTCTCTCTGGGACCAACAGAGTGCAGCGGAAACAAAGAGGTGGACAGGAGGC 417
Db 6 TTCTTCTCTCTGGGACCAACAGAGTGCAGCGGAAACAAAGAGGTGGACAGGAGGC 65
Qy 418 CAGGACACAGCGGGACATTTATCCAGAGGATTTCTAGAGCTTATTACAATCTGACCTG 477
Db 66 CAGGACACAGCGGGACATTTATCCAGAGGATTTCTAGAGCTTATTACAATCTGACCTG 125
Qy 478 AAGACCATGATGGGCATAGAAATGGGTCCATCGCTTTTGTCTCAGCGCGGCTTGTGATG 537
Db 126 AAGACCATGATGGGCATAGAAATGGGTCCATCGCTTTTGTCTCAGCGCGGCTTGTGATG 185
Qy 538 AAAACAGACTCAGACATGTTTCAATCAATGTTGAACTCAATGAGTTTCCCATCAGGAG 597
Db 186 AAAACAGACTCAGACATGTTTCAATCAATGTTGAACTCAATGAGTTTCCCATCAGGAG 245
Qy 598 AACAGAACACAGGTTTTTCACTGGCTTCTTGAACCTCAATGAGTTTCCCATCAGGAG 657
Db 246 AACAGAACACAGGTTTTTCACTGGCTTCTTGAACCTCAATGAGTTTCCCATCAGGAG 305
Qy 658 CCATTCAGCAAGTGGTTGTGAGTAAATCTGAATATCCGTGGGACAGGTACCCACCATTC 717
Db 306 CCATTCAGCAAGTGGTTGTGAGTAAATCTGAATATCCGTGGGACAGGTACCCACCATTC 365

Qy 718 TGCTCCGGCACCAGCTACGTGTTTCTGGGACGTCGGAGTCAGGTGTAACAATGTCCTC 777
 Db 366 TGCTCCGGCACCAGCTACGTGTTTCTGGGACGTCGGAGTCAGGTGTAACAATGTCCTC 425
 Qy 778 AAGAGCGTCCCATACATTAATACTGGAAGACGTGTTTGTGGGGCTCTGCTCGAAAGGCTG 837
 Db 426 AAGAGCGTCCCATACATTAATACTGGAAGACGTGTTTGTGGGGCTCTGCTCGAAAGGCTG 485
 Qy 838 AACATCAGATTGGAGGAGCTCCATCCAGCCGACCTTTTTCAGGGGGCTTACGCTTC 897
 Db 486 AACATCAGATTGGAGGAGCTCCATCCAGCCGACCTTTTTCAGGGGGCTTACGCTTC 545
 Qy 898 TCCGTATGCTCTTCAGAGGAGTCGTGGCTGCCACTTCATCAAGCCCTCGGACTCTCTTG 957
 Db 546 TCCGTATGCTCTTCAGAGGAGTCGTGGCTGCCACTTCATCAAGCCCTCGGACTCTCTTG 605
 Qy 958 GACTACTGGCAGGCTCTAGAGAAATCCCGGGGGGAAGATTGTCGCG 1003
 Db 606 GACTACTGGCAGGCTCTAGAGAAATCCCGGGGGGAAGATTGTCGCGC 651

RESULT 5
 US-10-027-632-21208
 ; Sequence 21208, Application US/10027632
 ; Publication No. US20030204075A9
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 ; FILE REFERENCE: 108827.129
 ; CURRENT APPLICATION NUMBER: US/10/027,632
 ; CURRENT FILING DATE: 2002-04-30
 ; PRIOR APPLICATION NUMBER: US 60/218,006
 ; PRIOR FILING DATE: 2000-07-12
 ; PRIOR APPLICATION NUMBER: US 60/198,676
 ; PRIOR FILING DATE: 2000-04-20
 ; PRIOR APPLICATION NUMBER: US 60/193,483
 ; PRIOR FILING DATE: 2000-03-29
 ; PRIOR APPLICATION NUMBER: US 60/185,218
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: US 60/167,363
 ; PRIOR FILING DATE: 1999-11-23
 ; PRIOR APPLICATION NUMBER: US 60/156,358
 ; PRIOR FILING DATE: 1999-09-28
 ; PRIOR APPLICATION NUMBER: US 60/146,002
 ; PRIOR FILING DATE: 1999-08-09
 ; NUMBER OF SEQ ID NOS: 325720
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 21208
 ; LENGTH: 681
 ; TYPE: DNA
 ; ORGANISM: Human
 US-10-027-632-21208

Query Match 63.3%; Score 640; DB 17; Length 681;
 Best Local Similarity 99.1%; Pred. No. 1.2e-201;
 Matches 640; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 Qy 358 TTCTTCCTCTCTGGGACCAACAGAGTCAGCGGAAACAAAGAGGTGGACGAGAGC 417
 Db 6 TTCTTCCTCTCTGGGACCAACAGAGTCAGCGGAAACAAAGAGGTGGACGAGAGC 65
 Qy 418 CAGCGACACGGGACATTAATCCAGAGGATTCCTAGACGTCTATTACAATCTGACCTG 477
 Db 66 CAGCGACACGGGACATTAATCCAGAGGATTCCTAGACGTCTATTACAATCTGACCTG 125
 Qy 478 AAGACCATGATGGGCATAGATGGGTCCATCGCTTTTGTCTCAGCGGGCTTTGTGATG 537
 Db 126 AAGACCATGATGGGCATAGATGGGTCCATCGCTTTTGTCTCAGCGGGCTTTGTGATG 185
 Qy 538 AAAACAGACTCAGACATGTTTCATCAATGTCGACTATCTGACTGAACTGCTTCTGAAGAA 597

Db 186 AAAACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGAACTGCTTCTGAAGAA 245
 Qy 598 AACAGACACACGAGTGTTCCTGCTGCTTCTGAACTCAATGAGTTTCCCATCAGGACG 657
 Db 246 AACAGAAACAACGAGTGTTCCTGCTGCTTCTGAACTCAATGAGTTTCCCATCAGGACG 305
 Qy 658 CCATTTCAGCAAGTGGTGTTCAGTAAATCTGAATATCCGTGGGACAGGTACCCACCATTC 717
 Db 306 CCATTTCAGCAAGTGGTGTTCAGTAAATCTGAATATCCGTGGGACAGGTACCCACCATTC 365
 Qy 718 TGCTCCGCAACCGGCTACGTTTTCCTGGGAGTCGTGGGAGTCAGGTGTAACAATGTCCTC 777
 Db 366 TGCTCCGCAACCGGCTACGTTTTCCTGGGAGTCGTGGGAGTCAGGTGTAACAATGTCCTC 425
 Qy 778 AAGAGCGTCCCATACATTAACCTGGAAGACGTGTTTGTGGGGCTCTGCTCGAAAGGCTG 837
 Db 426 AAGAGCGTCCCATACATTAACCTGGAAGACGTGTTTGTGGGGCTCTGCTCGAAAGGCTG 485
 Qy 838 AACATCAGATTGGAGGAGCTCCACTCCAGCCGACCTTTTTCAGGGGGCTTACGCTTC 897
 Db 486 AACATCAGATTGGAGGAGCTCCACTCCAGCCGACCTTTTTCAGGGGGCTTACGCTTC 545
 Qy 898 TCCGTATGCTCTTCAGAGGAGTCGTGGCTGCCACTTCATCAAGCCCTCGGACTCTCTTG 957
 Db 546 TCCGTATGCTCTTCAGAGGAGTCGTGGCTGCCACTTCATCAAGCCCTCGGACTCTCTTG 605
 Qy 958 GACTACTGGCAGGCTCTAGAGAAATCCCGGGGGGAAGATTGTCGCG 1003
 Db 606 GACTACTGGCAGGCTCTAGAGAAATCCCGGGGGGAAGATTGTCGCGC 651

RESULT 6
 US-10-027-632-21209
 ; Sequence 21209, Application US/10027632
 ; Publication No. US20030204075A9
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 ; FILE REFERENCE: 108827.129
 ; CURRENT APPLICATION NUMBER: US/10/027,632
 ; CURRENT FILING DATE: 2002-04-30
 ; PRIOR APPLICATION NUMBER: US 60/218,006
 ; PRIOR FILING DATE: 2000-07-12
 ; PRIOR APPLICATION NUMBER: US 60/198,676
 ; PRIOR FILING DATE: 2000-04-20
 ; PRIOR APPLICATION NUMBER: US 60/193,483
 ; PRIOR FILING DATE: 2000-03-29
 ; PRIOR APPLICATION NUMBER: US 60/185,218
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: US 60/167,363
 ; PRIOR FILING DATE: 1999-11-23
 ; PRIOR APPLICATION NUMBER: US 60/156,358
 ; PRIOR FILING DATE: 1999-09-28
 ; PRIOR APPLICATION NUMBER: US 60/146,002
 ; PRIOR FILING DATE: 1999-08-09
 ; NUMBER OF SEQ ID NOS: 325720
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 21209
 ; LENGTH: 681
 ; TYPE: DNA
 ; ORGANISM: Human
 US-10-027-632-21209

Query Match 63.3%; Score 640; DB 17; Length 681;
 Best Local Similarity 99.1%; Pred. No. 1.2e-201;
 Matches 640; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 Qy 358 TTCTTCCTCTCTGGGACCAACAGAGTCAGCGGAAACAAAGAGGTGGACGAGAGC 417
 Db 6 TTCTTCCTCTCTGGGACCAACAGAGTCAGCGGAAACAAAGAGGTGGACGAGAGC 65
 Qy 418 CAGCGACACGGGACATTAATCCAGAGGATTCCTAGACGTCTATTACAATCTGACCTG 477


```

Db 66 CAGGACACGGGGACATATCAGAAAGGATTTCTAGACGCTATTATCAATCTGACCCCTG 125
Qy 478 AAGACCATGATGGGATAGAGTGGGTCCATCGCTTTTCTCTCAGGCGGGCTTTGTGATG 537
Db 126 AAGACCATGATGGGATAGAGTGGGTCCATCCTTTTCTCTCAGGCGGGCTTTGTGATG 185
Qy 538 AAAACAGACTCAGACATGTTTCAATGTTGATCTATCTGACTGATGAACTGCTTCTGAAGAA 597
Db 186 AAAACAGACTCAGACATGTTTCAATGTTGATCTATCTGACTGATGAACTGCTTCTGAAGAA 245
Qy 598 AACAGAACCAACAGGTTTTTCACTGGCTTCTGAACTCAATGATGTTTCCCATCAGGCAG 657
Db 246 AACAGAACCAACAGGTTTTTCACTGGCTTCTGAACTCAATGATGTTTCCCATCAGRCAG 305
Qy 658 CCATTTCAGCAAGTGGTTTGTCACTGAGTAAATCTGAATATCCGTGGGACAGGTACCCACCATTC 717
Db 306 CCATTTCAGCAAGTGGTTTGTCACTGAGTAAATCTGAATATCCGTGGGACAGGTACCCACCATTC 365
Qy 718 TGCTCCGGCACCGGCTACGTTTCTGCGGACGTCGTCGAGTCAAGGTTGACAAATGTTCTCC 777
Db 366 TGCTCCGGCACCGGCTACGTTTCTGCGGACGTCGTCGAGTCAAGGTTGACAAATGTTCTCC 425
Qy 778 AAGACGCTCCCATACATTAATCTGAAGACGTTGTTGTGGGGCTCTGCTCGAAAGGCTG 837
Db 426 AAGACGCTCCCATACATTAATCTGAAGACGTTGTTGTGGGGCTCTGCTCGAAAGGCTG 485
Qy 838 AACATCAGATTGGAGGAGCTCACTCCAGCGGACCTTTTTCAGGGGGCTTACGCTTC 897
Db 486 AACATCAGATTGGAGGAGCTCACTCCAGCGGACCTTTTTCAGGGGGCTTACGCTTC 545
Qy 958 TCCGTATGCTCTTTTTCAGGAGGATCGTGGCTCGCACTTTCATCAAGCCCTCGGACTCTCTTG 957
Db 546 TCCGTATGCTCTTTTTCAGGAGGATCGTGGCTCGCACTTTCATCAAGCCCTCGGACTCTCTTG 605
Qy 958 GACTACTGGCAGGCTCTAGAGAAATTCGCGGGGGAAGATTGTCGCG 1003
Db 606 GACTACTGGCAGGCTCTAGAGAAATTCGCGGGGGAAGATTGTCGCGC 651

```

RESULT 7

```

US-10-027-632-21210
; Sequence 21210, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21210
; LENGTH: 681
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-21210

```

```

Query Match 63.3%; Score 640; DB 17; Length 681;
Best Local Similarity 99.1%; Pred. No. 1.2e-201;
Matches 640; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 358 TTCTTCTCTCTGGGGACCCACGACGATGTCAGCGAAACAAAGAGGTGGACCGAGGAGC 417
Db 6 TTCTTCTCTCTGGGGACCCACGACGATGTCAGCGAAACAAAGAGGTGGACCGAGGAGC 65
Qy 418 CAGGACACCGGGGACATTTATCCAGAAGGATTTCTTAGAAGTCTTATCAATCTGACCCCTG 477
Db 66 CAGGACACCGGGGACATTTATCCAGAAGGATTTCTTAGAAGTCTTATCAATCTGACCCCTG 125
Qy 478 AAGACCATGATGGGATAGAGTGGGTCCATCGCTTTTCTCTCAGGCGGGCTTTGTGATG 537
Db 126 AAGACCATGATGGGATAGAGTGGGTCCATCCTTTTCTCTCAGGCGGGCTTTGTGATG 185
Qy 538 AAAACAGACTCAGACATGTTTCAATGTTGATCTATCTGACTGATGAACTGCTTCTGAAGAA 597
Db 186 AAAACAGACTCAGACATGTTTCAATGTTGATCTATCTGACTGATGAACTGCTTCTGAAGAA 245
Qy 598 AACAGAACCAACAGGTTTTTCACTGGCTTCTGAACTCAATGATGTTTCCCATCAGGCAG 657
Db 246 AACAGAACCAACAGGTTTTTCACTGGCTTCTGAACTCAATGATGTTTCCCATCAGRCAG 305
Qy 658 CCATTTCAGCAAGTGGTTTGTCACTGAGTAAATCTGAATATCCGTGGGACAGGTACCCACCATTC 717
Db 306 CCATTTCAGCAAGTGGTTTGTCACTGAGTAAATCTGAATATCCGTGGGACAGGTACCCACCATTC 365
Qy 718 TGCTCCGGCACCGGCTACGTTTCTGCGGACGTCGTCGAGTCAAGGTTGACAAATGTTCTCC 777
Db 366 TGCTCCGGCACCGGCTACGTTTCTGCGGACGTCGTCGAGTCAAGGTTGACAAATGTTCTCC 425
Qy 778 AAGACGCTCCCATACATTAATCTGAAGACGTTGTTGTGGGGCTCTGCTCGAAAGGCTG 837
Db 426 AAGACGCTCCCATACATTAATCTGAAGACGTTGTTGTGGGGCTCTGCTCGAAAGGCTG 485
Qy 838 AACATCAGATTGGAGGAGCTCACTCCAGCGGACCTTTTTCAGGGGGCTTACGCTTC 897
Db 486 AACATCAGATTGGAGGAGCTCACTCCAGCGGACCTTTTTCAGGGGGCTTACGCTTC 545
Qy 898 TCCGTATGCTCTTTTTCAGGAGGATCGTGGCTCGCACTTTCATCAAGCCCTCGGACTCTCTTG 957
Db 546 TCCGTATGCTCTTTTTCAGGAGGATCGTGGCTCGCACTTTCATCAAGCCCTCGGACTCTCTTG 605
Qy 958 GACTACTGGCAGGCTCTAGAGAAATTCGCGGGGGAAGATTGTCGCG 1003
Db 606 GACTACTGGCAGGCTCTAGAGAAATTCGCGGGGGAAGATTGTCGCGC 651

```

RESULT 8

```

US-10-027-632-106210
; Sequence 106210, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002

```

```

; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 106210
; LENGTH: 516
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-106210

```

Query Match	51.0%;	Score	515.6;	DB	13;	Length	516;
Best Local Similarity	99.8%;	Pred. No.	2.6e-160;				
Matches	515;	Conservative	1;	Mismatches	0;	Indels	0;
Qy	350	TGAAGACATTCTTCCTCCTCGTGGGACCA	CCAGCAGCTGCAGCGGAAACAAAGAGCTTGAC	CC	409		
Db	1	TGAAGACATTCTTCCTCCTCGTGGGACCA	CCAGCAGCTGCAGCGGAAACAAAGAGCTTGAC	CC	60		
Qy	410	AGGAGACCCAGCGACACACGGGACATTAT	CCAGAAAGGATTTCTTAGAGCTCTATTACAATC	469			
Db	61	AGGAGACCCAGCGACACGGGACATTAT	CCAGAAAGGATTTCTTAGAGCTCTATTACAATC	120			
Qy	470	TGACCTTGAAGACCATGATGGGCATAGAAT	TGGGTCCATTCGCTTTTGTCTCAGCGGGCGT	529			
Db	121	TGACCTTGAAGACCATGATGGGCATAGAAT	TGGGTCCATTCGCTTTTGTCTCAGCGGGCGT	180			
Qy	530	TTGTGATGAAAAACAGACTTCAGACATGTT	CATCAATGTTGACTATCTGACTGAACCTGCTC	589			
Db	181	TTGTGATGAAAAACAGACTTCAGACATGTT	CATCAATGTTGACTATCTGACTGAACCTGCTC	240			
Qy	590	TGAAGAAAAACAGAAACACCAGGTTTTTTC	ACTGGCTCTTCTTGAAACTCAAATGAGTTTCCCA	649			
Db	241	TGAAGAAAAACAGAAACACCAGGTTTTTTC	ACTGGCTCTTCTTGAAACTCAAATGAGTTTCCCA	300			
Qy	650	TCAGGCAGCCATTTCAGCAAGTGGTTGTCT	AGTAAATCTGTAATATCCGTGGGACAGGTACC	709			
Db	301	TCAGGCAGCCATTTCAGCAAGTGGTTGTCT	AGTAAATCTGTAATATCCGTGGGACAGGTACC	360			
Qy	710	CACCATCTCTCGTCCGGACCGGCTACGTTT	CTCTGGCGAGCTGGCGAGTCAGGTGTACA	769			
Db	361	CACCATCTCTCGTCCGGACCGGCTACGTTT	CTCTGGCGAGCTGGCGAGTCAGGTGTACA	420			
Qy	770	ATGTCTCCAAGAGCGTCCCATACATAATTA	AACTGGGAAGACGTGTTTGTGGGGCTCTGCTCCTG	829			
Db	421	ATGTCTCCAAGAGCGTCCCATACATAATTA	AACTGGGAAGACGTGTTTGTGGGGCTCTGCTCCTG	480			
Qy	830	AAAGGCTGAACATCAGATTGGAGAGCTTCCA	CTCCC	865			
Db	481	AAAGGCTGAACATCAGATTGGAGAGCTTCCA	CTCCC	516			

```

RESULT 9
US-10-027-632-106210
; Sequence 106210, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.1129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28

```

```

; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 106210
; LENGTH: 516
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-106210

```

Query Match	51.0%;	Score 515.6;	DB 17;	Length 516;
Best Local Similarity	99.8;	Pred. No. 2.6e-160;		
Matches 515;	Conservative 1;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	350	TGAAGACATTCTTCCCTCTGGGACACACAGCAGCTGCAGCGGAAACAAAGAGGTGGACC	409	
Db	1	TGAAGACATTCTTCCCTCTGGGACACACAGCAGCTGCAGCGGAAACAAAGAGGTGGACC	60	
Qy	410	AGGAGAGCCACAGCGACACACGGGGAATTATTCAGAAAGGATTTCTTAGACGTCTATTACAATC	469	
Db	61	AGGAGAGCCACAGCGACACACGGGGAATTATTCAGAAAGGATTTCTTAGACGTCTATTACAATC	120	
Qy	470	TGACCTTGAAGACCATGATGGGCGATGAATGGGTCCATCGCTTTTGTCTCAGCGGGCGT	529	
Db	121	TGACCTTGAAGACCATGATGGGCGATGAATGGGTCCATCGCTTTTGTCTCAGCGGGCGT	180	
Qy	530	TTGTGATGAAAAACAGACTTCAGACATGTGTTCAATCAATGTTTGACTATCTGACTCAACTGCTTC	589	
Db	181	TTGTGATGAAAAACAGACTTCAGACATGTGTTCAATCAATGTTTGACTATCTGACTCAACTGCTTC	240	
Qy	590	TGAAGAAAAACAGAAACACAGGTTTTTTCATCTGGCTCTTGTGAAACTCAATGAGTTTCCCA	649	
Db	241	TGAAGAAAAACAGAAACACAGGTTTTTTCATCTGGCTCTTGTGAAACTCAATGAGTTTCCCA	300	
Qy	650	TCAGGAGCCATTTCAGCAAGTGGTTGTGCTAGTAAATCTGAAATATCCGTGGGACAGGTACC	709	
Db	301	TCAGGAGCCATTTCAGCAAGTGGTTGTGCTAGTAAATCTGAAATATCCGTGGGACAGGTACC	360	
Qy	710	CACCATTTCTCTCCGGACCGGCTACGTGTTTTCTGGCGAGCTGGCAGTCAAGTGTATCA	769	
Db	361	CACCATTTCTCTCCGGACCGGCTACGTGTTTTCTGGCGAGCTGGCAGTCAAGTGTATCA	420	
Qy	770	ATGTCTCCAAGAGCGTCCCATACATTAAACTGGGAAGACGTGTTTGTGGGGCTCTGCCTCG	829	
Db	421	ATGTCTCCAAGAGCGTCCCATACATTAAACTGGGAAGACGTGTTTGTGGGGCTCTGCCTCG	480	
Qy	830	AAAGGCTGAACATCAGATTGGAGGAGTCCACTCCC	865	
Db	481	AAAGGCTGAACATCAGATTGGAGGAGTCCACTCCC	516	

```

RESULT 10
US-10-027-632-106208
; Sequence 106208, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358

```

RESULT 12
US-10-027-632-106208
/ Sequence 106208, Application US/10027632
/ Publication No. US20030204075A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ CURRENT FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363

```
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 106208
; LENGTH: 516
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-106208
```

```
Query Match      50.8%; Score 514; DB 17; Length 516;
Best Local Similarity 99.6%; Pred. No. 9e-160;
Matches 514; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 350 TGAAGACATTCTTCTCTCTGGGGACCAACAGCAGTGCAGCGGAAACAAAGAGGTGGACC 409
Db 1 TGAAGACATTCTTCTCTCTGGGGACCAACAGCAGTGCAGCGGAAACAAAGAGGTGGACC 60

Qy 410 AGGAGAGCCAGCAGCAGCGGACATATCCAGAGGATTTCTAGACGTCTATTACAAATC 469
Db 61 AGGAGAGCCAGCAGCAGCGGACATATCCAGAGGATTTCTAGACGTCTATTACAAATC 120

Qy 470 TGACCCCTGAAGACCATGATGGGCATAGATGGGTCCATCGCTTTTGTCTCAGGGGGGT 529
Db 121 TGACCCCTGAAGACCATGATGGGCATAGATGGGTCCATCGCTTTTGTCTCAGGGGGGT 180

Qy 530 TTGTGATGAAACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGAACTGCTTC 589
Db 181 TTGTGATGAAACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGAACTGCTTC 240

Qy 590 TGAAGAAAAACAGAACACAGGTTTTTCATCGGCTTCTGAAACTCAATGAGTTTCCCA 649
Db 241 TGAAGAAAAACAGAACACAGGTTTTTCATCGGCTTCTGAAACTCAATGAGTTTCCCA 300

Qy 650 TCAGGACCCATTTCAGCAAGTGGTTGTGAGTAATCTGAATATCCGTGGGACAGGTACC 709
Db 301 TCAGGACCCATTTCAGCAAGTGGTTGTGAGTAATCTGAATATCCGTGGGACAGGTACC 360

Qy 710 CACCATTCCTGCTCCGGCACCGGCTACGTTTCTGCGGACGTGGCGAGTCAGGTGTACA 769
Db 361 CACCATTCCTGCTCCGGCACCGGCTACGTTTCTGCGGACGTGGCGAGTCAGGTGTACA 420

Qy 770 ATGTCTCAAGAGCGTCCCATACATTAACCTGGAAGACGTGTTGTGGGGCTCTGCCTCG 829
Db 421 ATGTCTCAAGAGCGTCCCATACATTAACCTGGAAGACGTGTTGTGGGGCTCTGCCTCG 480

Qy 830 AAAGGCTGAACATCAGATTGGAGGAGCTCCACTCCC 865
Db 481 AAAGGCTGAACATCAGATTGGAGGAGCTCCACTCCC 516
```

```
RESULT 13
US-10-027-632-106209
; Sequence 106209, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
```

```
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 106209
; LENGTH: 516
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-106209
```

```
Query Match      50.8%; Score 514; DB 17; Length 516;
Best Local Similarity 99.6%; Pred. No. 9e-160;
Matches 514; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 350 TGAAGACATTCTTCTCTCTGGGGACCAACAGCAGTGCAGCGGAAACAAAGAGGTGGACC 409
Db 1 TGAAGACATTCTTCTCTCTGGGGACCAACAGCAGTGCAGCGGAAACAAAGAGGTGGACC 60

Qy 410 AGGAGAGCCAGCAGCAGCGGACATATCCAGAGGATTTCTAGACGTCTATTACAAATC 469
Db 61 AGGAGAGCCAGCAGCAGCGGACATATCCAGAGGATTTCTAGACGTCTATTACAAATC 120

Qy 470 TGACCCCTGAAGACCATGATGGGCATAGATGGGTCCATCGCTTTTGTCTCAGGGGGGT 529
Db 121 TGACCCCTGAAGACCATGATGGGCATAGATGGGTCCATCGCTTTTGTCTCAGGGGGGT 180

Qy 530 TTGTGATGAAACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGAACTGCTTC 589
Db 181 TTGTGATGAAACAGACTCAGACATGTTTCATCAATGTTGACTATCTGACTGAACTGCTTC 240

Qy 590 TGAAGAAAAACAGAACACAGGTTTTTCATCGGCTTCTTGAACACTCAATGAGTTTCCCA 649
Db 241 TGAAGAAAAACAGAACACAGGTTTTTCATCGGCTTCTTGAACACTCAATGAGTTTCCCA 300

Qy 650 TCAGGACCCATTTCAGCAAGTGGTTGTGAGTAATCTGAATATCCGTGGGACAGGTACC 709
Db 301 TCAGGACCCATTTCAGCAAGTGGTTGTGAGTAATCTGAATATCCGTGGGACAGGTACC 360

Qy 710 CACCATTCCTGCTCCGGCACCGGCTACGTTTCTGCGGACGTGGCGAGTCAGGTGTACA 769
Db 361 CACCATTCCTGCTCCGGCACCGGCTACGTTTCTGCGGACGTGGCGAGTCAGGTGTACA 420

Qy 770 ATGTCTCAAGAGCGTCCCATACATTAACCTGGAAGACGTGTTGTGGGGCTCTGCCTCG 829
Db 421 ATGTCTCAAGAGCGTCCCATACATTAACCTGGAAGACGTGTTGTGGGGCTCTGCCTCG 480

Qy 830 AAAGGCTGAACATCAGATTGGAGGAGCTCCACTCCC 865
Db 481 AAAGGCTGAACATCAGATTGGAGGAGCTCCACTCCC 516
```

```
RESULT 14
US-10-029-386-13121
; Sequence 13121, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GI
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: ABOICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 13121
; LENGTH: 596
; TYPE: DNA
```

Twining

Db 542 AATTCTTATCAACGAGCCCAATAAATGTGAGAAAACATTCCTTTCTGTTATCCTCA 601

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2005, 05:52:47 ; Search time 603.053 Seconds
(without alignments)
9924.272 Million cell updates/sec

Title: US-10-777-828-8

Perfect score: 1011

Sequence: 1 ccacctcagcctctagcat.....aagattgcgcctgtctga 1011

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1011	100.0	1011	3	AAA27959 Human bet
2	1006.2	99.5	3409	10	ABZ56919 B3GALT nu
3	1006.2	99.5	10562	3	AAA93876 Human bet
4	928.8	91.9	933	10	ABZ56916 B3GALT nu
5	928.6	91.8	2629	10	ABZ56918 B3GALT nu
6	928.6	91.8	2762	10	ABZ56917 B3GALT nu
7	928.6	91.8	2775	3	AAA93875 Human bet
8	928.2	91.8	933	12	ADK68456 Human bet
9	405.2	40.1	596	12	ACH79926 Human gen
10	163.4	16.2	1037	10	ABZ56909 B3GALT nu
11	163.4	16.2	1739	2	AAQ67067 Beta-1,3-
12	161.8	16.0	2168	10	ABZ56910 B3GALT nu
13	133.6	13.2	1092	12	ADQ00390 Novel hum
14	133.6	13.2	1092	12	ADN98821 Novel hum
15	133.6	13.2	1358	10	ADP76806 Novel hum
16	133.6	13.2	1773	2	AAx87193 Human Den
17	133.6	13.2	1897	6	ABK51201 Human CDN
18	133.6	13.2	1897	12	ADO19801 Human PRO
19	133.6	13.2	2095	3	Az65022 Membrane-
20	133.6	13.2	2095	4	AAs46005 Human DNA

ALIGNMENTS

RESULT 1
AAA27959
ID AAA27959 standard; DNA; 1011 BP.
XX
AC AAA27959;
XX
DT 15-AUG-2000 (first entry)
XX
DE Human beta3Gal-T5 gene sequence.
XX
KW UDP-D-galactose; beta-N-acetylglucosamine beta1,3-galactosyltransferase;
KW Beta3Gal-T5; ss; human; chromosome 21q22.3; galactosylation;
KW beta1,3-galactosyl glycosylated saccharide production; glycopeptide;
KW glycoprotein.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 79..1011
FT /*tag= a
FT /*product= "Beta3Gal-T5"
FT primer_bind complement(79..98)
FT /*tag= b
FT primer_bind complement(150..170)
FT /*tag= c
FT primer_bind 991..1011
FT /*tag= d
PN WO200029558-A1.
XX
PD 25-MAY-2000.
XX
PF 11-NOV-1999; 99WO-US026807.
XX
PR 13-NOV-1998; 98DK-00001483.
XX (CLAU/) CLAUSEN H.
XX Clausen H, Amado M;
XX WPI; 2000-399728/34.
XX P-PSDB; AAY94641.
XX

PT Novel nucleic acid sequence encoding human UDP-galactose:beta-N-
PT acetylglucosamine beta1,3-galactosyltransferase useful for obtaining beta
PT 1,3-galactosyl glycosylated saccharides and glycopeptides or
PT glycoproteins.

PS Claim 7; Fig 1; 74pp; English.

XX The present invention relates to a nucleic acid sequence encoding UDP-D-
XX galactose:beta-N-acetylglucosamine beta1,3-galactosyltransferase
CC (beta3Gal-T5). Beta3 transferases add galactose to the hydroxy group at
CC carbon 3 of 2-acetamido-2-deoxy-D-glucose (GlcNAc). The present sequence
CC represents the human beta3Gal-T5 gene sequence. The beta3Gal-T5 gene is
CC located on human chromosome 21q22.3. Beta3Gal-T5 is a type II
CC transmembrane glycoprotein. The invention also relates to the beta3Gal-T5
CC protein sequence, a nucleic acid vector comprising the beta3Gal-T5
CC nucleotide sequence, a host cell comprising the vector, and a method for
CC the production of the beta3Gal-T5 protein from the host cells. The
CC methods of the invention can be used for recombinant production of
CC beta3Gal-T5 for use as a catalyst and for recombinant production of
CC peptides or proteins with appropriate galactosylation. The beta3Gal-T5
CC protein can be used to obtain beta1,3-galactosyl glycosylated
CC saccharides, glycopeptides or glycoproteins
XX
SQ Sequence 1011 BP; 247 A; 256 C; 251 G; 257 T; 0 U; 0 Other;

Query Match 100.0%; Score 1011; DB 3; Length 1011;
Best Local Similarity 100.0%; Pred. No. 5e-304;
Matches 1011; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCACCTCAGCTCCTAGCATAAACATAGACATCCTCATGCTTTGAGGTCTAATCAT 60
DB 1 CCACCTCAGCTCCTAGCATAAACATAGACATCCTCATGCTTTGAGGTCTAATCAT 60
QY 61 GGATTTTGTCTTTCAGATGGCTTCCGAGATGAGATTGATATATCTGCTTCTG 120
DB 61 GGATTTTGTCTTTCAGATGGCTTCCGAGATGAGATTGATATATCTGCTTCTG 120
QY 121 GTTCGGGGCTCTTTGTTTATTTAGCATGTACAGTCTAAATCTCTTCAAGAACAG 180
DB 121 GTTCGGGGCTCTTTGTTTATTTAGCATGTACAGTCTAAATCTCTTCAAGAACAG 180
QY 181 TCCTTTGTTTACAGAAAGCGGAATCTCTTAAGCTCCAGATACAGACTGCAGGCAG 240
DB 181 TCCTTTGTTTACAGAAAGCGGAATCTCTTAAGCTCCAGATACAGACTGCAGGCAG 240
QY 241 ACACCTCCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
DB 241 ACACCTCCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
QY 301 GCCATCCGGCAGACGTGGGGGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGATTC 360
DB 301 GCCATCCGGCAGACGTGGGGGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGATTC 360
QY 361 TTCTCTCTGGGGACACACAGCAGTGCAGCGAAACAAAGAGAGTGCAGCAGAGCCAG 420
DB 361 TTCTCTCTGGGGACACACAGCAGTGCAGCGAAACAAAGAGAGTGCAGCAGAGCCAG 420
QY 421 CGACACGGGGACATPATTCAGAGAGATTTCCTAGACGTCTATTACAATCTGACCTGAAG 480
DB 421 CGACACGGGGACATPATTCAGAGAGATTTCCTAGACGTCTATTACAATCTGACCTGAAG 480
QY 481 ACCATGATGGGCATAGATGGTCCATCGCTTTGCTCAGGGCGCTTTGTGATGAA 540
DB 481 ACCATGATGGGCATAGATGGTCCATCGCTTTGCTCAGGGCGCTTTGTGATGAA 540
QY 541 ACAGACTCAGACATGCTTCATCAATGTTGACTATCTGACTGAACTGCTTCTGAAGAAAAC 600
DB 541 ACAGACTCAGACATGCTTCATCAATGTTGACTATCTGACTGAACTGCTTCTGAAGAAAAC 600
QY 601 AGAACACACAGGTTTTCTCAGTGGCTTCTTGAACTCAATGATGTTCCCATCAGGCAGCA 660
DB 601 AGAACACACAGGTTTTCTCAGTGGCTTCTTGAACTCAATGATGTTCCCATCAGGCAGCA 660

QY 661 TTCAGCAAGTGGTTTGTCTAGTAATCTGAATATCCGTGGGACAGGTACCCACCATTTCTGC 720
DB 661 TTCAGCAAGTGGTTTGTCTAGTAATCTGAATATCCGTGGGACAGGTACCCACCATTTCTGC 720
QY 721 TCCGGCACCCGGCTACGTGTTTTCTGCGCAGCTGGCAGTCAAGTGTACAATGTCTCCAAG 780
DB 721 TCCGGCACCCGGCTACGTGTTTTCTGCGCAGCTGGCAGTCAAGTGTACAATGTCTCCAAG 780
QY 781 AGCGTCCCATATCAATTAACCTGGAAGAGCTGTTTGTGGGGCTCTGCTCGAAGGCTGAAC 840
DB 781 AGCGTCCCATATCAATTAACCTGGAAGAGCTGTTTGTGGGGCTCTGCTCGAAGGCTGAAC 840
QY 841 ATCAGATTGGAGGAGCTCCACTCCAGCCGACCTTTTTCAGGGGGCTTACGCTTCTCC 900
DB 841 ATCAGATTGGAGGAGCTCCACTCCAGCCGACCTTTTTCAGGGGGCTTACGCTTCTCC 900
QY 901 GTATGCTCTTTCAGGAGGATCGTGGCCTGCGACTTCATCAAGCCTCGGACTCTCTTGGAC 960
DB 901 GTATGCTCTTTCAGGAGGATCGTGGCCTGCGACTTCATCAAGCCTCGGACTCTCTTGGAC 960
QY 961 TACTGGCAGGCTCTAGAGAAATCCCGGGGGAGAGATTGTCGGCTGTCTGA 1011
DB 961 TACTGGCAGGCTCTAGAGAAATCCCGGGGGAGAGATTGTCGGCTGTCTGA 1011

RESULT 2
ABZ56919

ID ABZ56919 standard; DNA; 3409 BP.

XX ABZ56919;

DT 04-APR-2003 (first entry)

XX B3GALT nucleic acid sequence # SEQ ID 11.

XX B3GALT; beta-1,3 galactosyltransferase; p53; cancer; breast; colon;
KW kidney; lung; ovary; gene; ds.

XX Homo sapiens.

XX WO200299044-A2.

XX 12-DEC-2002.

XX 02-JUN-2002; 2002WO-US017356.

XX 05-JUN-2001; 2001US-0296076P.

XX 10-OCT-2001; 2001US-0328605P.

XX 15-FEB-2002; 2002US-0357253P.

XX (EXEL-) EXELIXIS INC.

XX Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;

XX WPI; 2003-156849/15.

XX Identifying p53 pathway modulating agents with B3GALT genes, useful for
PT the diagnosis and treatment of disorders associated with defects in the
PT p53 pathway, such as cancer of the breast, colon, kidneys, lung and
PT ovary.

XX Disclosure; Page 55-57; 82pp; English.

XX The invention relates to identifying a candidate p53 pathway modulating
CC agent in humans that is referred to in the specification as B3GALT (beta-
CC 1,3 galactosyltransferase). Also disclosed is a method for diagnosing a
CC disease in a patient, by contacting a sample with a probe for B3GALT
CC expression, and comparing the results with a control, and determining
CC whether the results indicate a likelihood of disease. Methods and
CC compositions of the invention are useful for the diagnosis and treatment
CC of disorders associated with defects in the p53 pathway, such as cancer
CC of the breast, colon, kidneys, lung and ovary. The current sequence
CC represents a B3GALT nucleic acid sequence referred to in the disclosure

CC of the invention
SQ Sequence 3409 BP; 833 A; 842 C; 833 G; 901 T; 0 U; 0 Other;
Query Match 99.5%; Score 1006.2; DB 10; Length 3409;
Best Local Similarity 99.7%; Pred. No. 3.2e-302;
Matches 1008; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CCACCTCAGCCTCTAGCATAAACTAGACACATCCTCATGCTTTTGAGGCTTAATCATTT 60
Db 1003 CCACCTCAGCCTCTAGCATAAACTAGACACATCCTCATGCTTTTGAGGCTTAATCATTT 1062
QY 61 GGATTTTGTCTTTTCAGATGCTTTCCGAAGATGAGATTGATGATATATCTGCCTTCTG 120
Db 1063 GGATTTTGTCTTTTCAGATGCTTTCCGAAGATGAGATTGATGATATATTTGCCCTTCTG 1122
QY 121 GTTCTGGGGGCTCTTTGTTGTTATTTTACATGATGATGATGATGATGATGATGATGAT 180
Db 1123 GTTCTGGGGGCTCTTTGTTGTTATTTTACATGATGATGATGATGATGATGATGATGAT 1182
QY 181 TCCTTTGTTTACAAGAAAGACGGGAATCTTCTTAAGCTCCAGATACAGACTGCAGGCAG 240
Db 1183 TCCTTTGTTTACAAGAAAGACGGGAATCTTCTTAAGCTCCAGATACAGACTGCAGGCAG 1242
QY 241 ACACCTCCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
Db 1243 ACACCTCCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1302
QY 301 GCCATCCGCGACAGCTGGGGGAAAGAGAGGAGCGGTGAAGGGAAGACAGCTGAAGCATTC 360
Db 1303 GCCATCCGCGACAGCTGGGGGAAAGAGAGGAGCGGTGAAGGGAAGACAGCTGAAGCATTC 1362
QY 361 TTCTCTCTGGGACACACAGCAGTGCAGCGGAAACAAAAGAGGTGGACACAGAGAGCCAG 420
Db 1363 TTCTCTCTGGGACACACAGCAGTGCAGCGGAAACAAAAGAGGTGGACACAGAGAGCCAG 1422
QY 421 CGACACGGGGACATTTATCAGAAAGATTTCTAGACGCTCTATTACAATCTGACCCCTGAAG 480
Db 1423 CGACACGGGGACATTTATCAGAAAGATTTCTAGACGCTCTATTACAATCTGACCCCTGAAG 1482
QY 481 ACCATGATGGCAGATAGATGGTTCATCGCTTTGCTCTCAGGGGGCTTTGTGATGAAA 540
Db 1483 ACCATGATGGCAGATAGATGGTTCATCGCTTTGCTCTCAGGGGGCTTTGTGATGAAA 1542
QY 541 ACAGACTCAGACATGCTTCATCAATCTGACTATCTGACTGACTGCTTCTGAAGAAAAAC 600
Db 1543 ACAGACTCAGACATGCTTCATCAATCTGACTATCTGACTGACTGCTTCTGAAGAAAAAC 1602
QY 601 AGAACACACAGGTTTTTCACTGGCTTTCTTGAAACTCAATGAGTTTTCCCATCAGGCAGCCA 660
Db 1603 AGAACACACAGGTTTTTCACTGGCTTTCTTGAAACTCAATGAGTTTTCCCATCAGGCAGCCA 1662
QY 661 TTCAGCAAGTGGTTGTCAGTAAATCTGAATATCTCGTGGGACAGTATCCCAACCATCTGC 720
Db 1663 TTCAGCAAGTGGTTGTCAGTAAATCTGAATATCTCGTGGGACAGTATCCCAACCATCTGC 1722
QY 721 TCCGCGACCGGCTAGCTGCTTTCTCGGACAGTGGGAGTCAAGTGTCAATGTTCTTCCAAG 780
Db 1723 TCCGCGACCGGCTAGCTGCTTTCTCGGACAGTGGGAGTCAAGTGTCAATGTTCTTCCAAG 1782
QY 781 AGCGTCCCATATTAATTAAGAGACGTTGTTGTTGGGCTCTGCTCGAAAGGCTGAAC 840
Db 1783 AGCGTCCCATATTAATTAAGAGACGTTGTTGTTGGGCTCTGCTCGAAAGGCTGAAC 1842
QY 841 ATCAGATTGGAGGAGCTCCATCCCGAGCGGACCTTTTTCAGGGGGCTTACGCTTCTCC 900
Db 1843 ATCAGATTGGAGGAGCTCCATCCCGAGCGGACCTTTTTCAGGGGGCTTACGCTTCTCC 1902
QY 901 GTATGCTCTTCAGGAGATCGTGGCTGCCACTTCATCAAGCCCTCGGACTCTCTTGGAC 960
Db 1903 GTATGCTCTTCAGGAGATCGTGGCTGCCACTTCATCAAGCCCTCGGACTCTCTTGGAC 1962
QY 961 TACTGGCAGGCTCTAGAGAAATTCGCGGGGGAAGATTGTCCGCTGTCTGA 1011

Db 1963 TACTGGCAGGCTCTAGAGAAATTCGCGGGGGAAGATTGTCCGCTGTCTGA 2013
RESULT 3
AAA93876
ID AAA93876 standard; DNA; 10562 BP.
XX AAA93876;
XX AC
DT 15-JAN-2001 (first entry)
XX Human beta3Gal-T5 encoding DNA.
XX Beta-1,3 galactose transferase; treatment; diagnosis; cancer; human;
XX digestive system; beta3Gal-T5; ds.
XX Homo sapiens.
XX WO200050608-A1.
XX 31-AUG-2000.
XX 24-FEB-2000; 2000WO-JP001070.
XX 25-FEB-1999; 99JP-00047571.
XX (KYOW) KYOWA HAKKO KOGYO KK.
XX Narimatsu H, Ishiki S, Togayachi A, Sasaki K;
XX WPI; 2000-549409/50.
DR Beta-1,3 galactose transferase and DNA encoding it, useful for synthesis
XX of type 1 sialyl Lewis, a carbohydrate for treatment of digestive system
XX cancer.
PS Claim 31; Page 103-111; 123pp; Japanese.
CC This invention relates to a polypeptide (I) with beta-1,3 galactose
CC transferase activity, or variants of (I) comprising amino acid additions,
CC deletions and/or substitutions. Included in the invention is DNA encoding
CC all or part of (I); expression vectors containing the DNA, host cells
CC transformed by the vectors; a method for the preparation of the
CC polypeptide by culture of the transformants or by expression in the milk
CC of a transgenic mammal, and antibodies recognising (I). The Beta-1,3
CC galactose transferase protein transfers galactose by beta-1,3 bonding to
CC N-acetylglucosamine present in a non-cyclic carbohydrate chain (such as
CC GlcNAc) to give Galbeta1-3GlcNAc. The protein and DNA
CC encoding it are useful for the treatment and diagnosis of cancer of the
CC digestive system. The present sequence represents a Beta3Gal-T5 encoding
CC DNA sequence
SQ Sequence 10562 BP; 2610 A; 2415 C; 2574 G; 2963 T; 0 U; 0 Other;
Query Match 99.5%; Score 1006.2; DB 3; Length 10562;
Best Local Similarity 99.7%; Pred. No. 6.3e-302;
Matches 1008; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CCACCTCAGCCTCTAGCATAAACTAGACACATCCTCATGCTTTTGAGGCTTAATCATTT 60
Db 8156 CCACCTCAGCCTCTAGCATAAACTAGACACATCCTCATGCTTTTGAGGCTTAATCATTT 8215
QY 61 GGATTTTGTCTTTTCAGATGCTTTCCGAAGATGAGATTGATGATATATCTGCCTTCTG 120
Db 8216 GGATTTTGTCTTTTCAGATGCTTTCCGAAGATGAGATTGATGATATATTTGCCCTTCTG 8275
QY 121 GTTCTGGGGGCTCTTTGTTGTTATTTTACATGATGATGATGATGATGATGATGATGAT 180
Db 8276 GTTCTGGGGGCTCTTTGTTGTTATTTTACATGATGATGATGATGATGATGATGATGAT 8335
QY 181 TCCTTTGTTTACAAGAAAGACGGGAATCTTCTTAAGCTCCAGATACAGACTGCAGGCAG 240
Db 181 TCCTTTGTTTACAAGAAAGACGGGAATCTTCTTAAGCTCCAGATACAGACTGCAGGCAG 240

QY 619 ACTGGCTTCTTGAACACTCAATGAGTTTCCATCAGGAGCCATTCCAGCAAGTGGTTTGTG 678
 Db 541 ACTGGCTTCTTGAACACTCAATGAGTTTCCATCAGGAGCCATTCCAGCAAGTGGTTTGTG 600
 QY 679 AGTAATCTGATATCCGTGGGACAGGTACCCACCAATCTGCTCGGACCGGCTACGTG 738
 Db 601 AGTAATCTGATATCCGTGGGACAGGTACCCACCAATCTGCTCGGACCGGCTACGTG 660
 QY 739 TTTTCTGGCGAGCTGGCGAGTCAGGTGTACAAATGTCTCCAGAGCGTCCCATACATATAA 798
 Db 661 TTTTCTGGCGAGCTGGCGAGTCAGGTGTACAAATGTCTCCAGAGCGTCCCATACATATAA 720
 QY 799 CTGGAAGACGTGTTTGTGGGCTCTGCTCGAAGGCTGAACATCAGATTGGAGGAGCTC 858
 Db 721 CTGGAAGACGTGTTTGTGGGCTCTGCTCGAAGGCTGAACATCAGATTGGAGGAGCTC 780
 QY 859 CACTCCAGCCGACCTTTTTCAGGGGGCTTACGCTTCTCGGTATGCTCTTCAAGAGG 918
 Db 781 CACTCCAGCCGACCTTTTTCAGGGGGCTTACGCTTCTCGGTATGCTCTTCAAGAGG 840
 QY 919 ATCGTGGCTGCACTTCAATCAAGCTCGGACTCTCTTGGACTCTCTTGGACTCTGGAGGCTCTAGAG 978
 Db 841 ATCGTGGCTGCACTTCAATCAAGCTCGGACTCTCTTGGACTCTCTTGGACTCTGGAGGCTCTAGAG 900
 QY 979 AATTCCCGGGGGAAGATTGTCGGCTGTCTGA 1011
 Db 901 AATTCCCGGGGGAAGATTGTCGGCTGTCTGA 933

RESULT 5
 ABZ56918
 ID ABZ56918 standard; DNA; 2629 BP.
 XX AC ABZ56918;
 XX DT 04-APR-2003 (first entry)
 XX DE B3GALT nucleic acid sequence # SEQ ID 10.
 XX KW B3GALT; beta-1,3 galactosyltransferase; p53; cancer; breast; colon;
 XX KW kidney; lung; ovary; gene; ds.
 XX OS Homo sapiens.
 XX PN WO200299044-A2.
 XX PD 12-DEC-2002.
 XX PF 02-JUN-2002; 2002WO-US017356.
 XX PR 05-JUN-2001; 2001US-0296076P.
 XX PR 10-OCT-2001; 2001US-0328605P.
 XX PR 15-FEB-2002; 2002US-0357253P.
 XX PA (EXEL-) EXELIXIS INC.
 XX PI Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
 XX WPI; 2003-156949/15.
 XX Identifying p53 pathway modulating agents with B3GALT genes, useful for
 PT the diagnosis and treatment of disorders associated with defects in the
 PT p53 pathway, such as cancer of the breast, colon, kidneys, lung and
 PT ovary.
 XX Disclosure; Page 53-55; 82pp; English.
 XX The invention relates to identifying a candidate p53 pathway modulating
 CC agent in humans that is referred to in the specification as B3GALT (beta-
 CC 1,3 galactosyltransferase). Also disclosed is a method for diagnosing a
 CC disease in a patient, by contacting a sample with a probe for B3GALT
 CC expression, and comparing the results with a control, and determining

CC whether the results indicate a likelihood of disease. Methods and
 CC compositions of the invention are useful for the diagnosis and treatment
 CC of disorders associated with defects in the p53 pathway, such as cancer
 CC of the breast, colon, kidneys, lung and ovary. The current sequence
 CC represents a B3GALT nucleic acid sequence referred to in the disclosure
 CC of the invention
 XX SQ Sequence 2629 BP; 632 A; 665 C; 644 G; 688 T; 0 U; 0 Other;
 Query Match 91.8%; Score 928.6; DB 10; Length 2629;
 Best Local Similarity 99.6%; Pred. No. 4.4e-278;
 Matches 931; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 77 AGATGGCTTTCCCGAAGATGAGATTGATGTATATCTGCTTCTGTTCTGGGGGCTCTTT 136
 Db 299 AAATGGCTTTCCCGAAGATGAGATTGATGTATATTTGCTTCTGTTCTGGGGGCTCTTT 358
 QY 137 GTTTGTATTATAGCATGTACAGTCTAAATCTCTTCAAAAGAACAGTCTTTGTTTACAAGA 196
 Db 359 GTTTGTATTATAGCATGTACAGTCTAAATCTCTTCAAAAGAACAGTCTTTGTTTACAAGA 418
 QY 197 AAGACGGGAACCTTCTTAAAGCTCCAGATACAGACTGAGGAGACACCTCCCTTCTCTCG 256
 Db 419 AAGACGGGAACCTTCTTAAAGCTCCAGATACAGACTGAGGAGACACCTCCCTTCTCTCG 478
 QY 257 TCCTGCTGGTGACCTCATCCCAAAACAGTTTGGCTGAGCGATGGCCATCCGGCAGAGCGT 316
 Db 479 TCCTGCTGGTGACCTCATCCCAAAACAGTTTGGCTGAGCGATGGCCATCCGGCAGAGCGT 538
 QY 317 GGGGAAAGAGAGGACGGTGAAGGAAAGAGCTGAAGACATTTCTTCTCTCTGGGAGCA 376
 Db 539 GGGGAAAGAGAGGATGTTGAAGGAAAGAGCTGAAGACATTTCTTCTCTCTGGGAGCA 598
 QY 377 CCAGCAGTGACGCGAACAAGAGGTTGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 436
 Db 599 CCAGCAGTGACGCGAACAAGAGGTTGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 658
 QY 437 TCCAGAAGATTCTCTAGACGTCTATTACAATCTGACCTGAAAGACCATGATGGGCATAG 496
 Db 659 TCCAGAAGATTCTCTAGACGTCTATTACAATCTGACCTGAAAGACCATGATGGGCATAG 718
 QY 497 AATGGGTCATCGCTTTTGTCTCAGCGGGGTTTGTGATGAAAACAGACTCAGACATGT 556
 Db 719 AATGGGTCATCGCTTTTGTCTCAGCGGGGTTTGTGATGAAAACAGACTCAGACATGT 778
 QY 557 TCATCAATGTTGACTATCTGACTGAACCTGCTTCTGAAGAAAACAGAACACAGGTTT 616
 Db 779 TCATCAATGTTGACTATCTGACTGAACCTGCTTCTGAAGAAAACAGAACACAGGTTT 838
 QY 617 TCATCTGGCTTCTTGAAACTCAATGAGTTTCCCATCAGGAGCCCATTCAGCAAGTGGTTTG 676
 Db 839 TCATCTGGCTTCTTGAAACTCAATGAGTTTCCCATCAGGAGCCCATTCAGCAAGTGGTTTG 898
 QY 677 TCAGTAAATCTGAATATCCGTGGGACAGGTACCAACCATTTCTGCTCCGGACACCGGCTAG 736
 Db 899 TCAGTAAATCTGAATATCCGTGGGACAGGTACCAACCATTTCTGCTCCGGACACCGGCTAG 958
 QY 737 TGTCTTCTGGCGAGCTGGCGAGTCAGGTGTACATGTCTCCAAGAGCGTCCCATACATTA 796
 Db 959 TGTCTTCTGGCGAGCTGGCGAGTCAGGTGTACATGTCTCCAAGAGCGTCCCATACATTA 1018
 QY 797 AACTGGAAGAGCTGTTTGTGGGGCTCTGCTCGAAAGGCTGAACATCAGATTGGAGGAGC 856
 Db 1019 AACTGGAAGAGCTGTTTGTGGGGCTCTGCTCGAAAGGCTGAACATCAGATTGGAGGAGC 1078
 QY 857 TCCACTCCAGCCGACCTTTTTCAGGGGGCTTACGCTTCTCCGTATGCTCTTTCAGGA 916
 Db 1079 TCCACTCCAGCCGACCTTTTTCAGGGGGCTTACGCTTCTCCGTATGCTCTTTCAGGA 1138
 QY 917 GGATCGTGGCTGGCCTTCAATCAAGCTCCGACTCTTTCGACTTCTGAGGAGGCTCTAG 976
 Db 1139 GGATCGTGGCTGGCCTTCAATCAAGCTCCGACTCTTTCGACTTCTGAGGAGGCTCTAG 1198

QY 977 AGAATTCCTCCGGGGAAGATTGTCCGCTGTCTGA 1011
 DB 1199 AGAATTCCTCCGGGGAAGATTGTCCGCTGTCTGA 1233

RESULT 6

ABZ56917
 ID ABZ56917 standard; DNA; 2762 BP.

XX ABZ56917;

DT 04-APR-2003 (first entry)

XX B3GALT nucleic acid sequence # SEQ ID 9.

DE B3GALT; beta-1,3 galactosyltransferase; p53; cancer; breast; colon;
 KW kidney; lung; ovary; gene; ds.

XX Homo sapiens.

XX WO200299044-A2.

PN 12-DEC-2002.

XX 02-JUN-2002; 2002WO-US017356.

XX 05-JUN-2001; 2001US-0296078P.

PR 10-OCT-2001; 2001US-0328605P.

PR 15-FEB-2002; 2002US-0357253P.

XX (EXEL-) EXELIXIS INC.

XX Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;

PI WPI; 2003-156849/15.

DR Identifying p53 pathway modulating agents with B3GALT genes, useful for
 PT the diagnosis and treatment of disorders associated with defects in the
 PT p53 pathway, such as cancer of the breast, colon, kidneys, lung and
 PT ovary.

XX Example 5; Page 52-53; 82pp; English.

XX The invention relates to identifying a candidate p53 pathway modulating
 CC agent in humans that is referred to in the specification as B3GALT (beta-
 CC 1,3 galactosyltransferase). Also disclosed is a method for diagnosing a
 CC disease in a patient, by contacting a sample with a probe for B3GALT
 CC expression, and comparing the results with a control, and determining
 CC whether the results indicate a likelihood of disease. Methods and
 CC compositions of the invention are useful for the diagnosis and treatment
 CC of disorders associated with defects in the p53 pathway, such as cancer
 CC of the breast, colon, kidneys, lung and ovary. The current sequence
 CC represents a B3GALT nucleic acid sequence referred to in an example from
 CC the invention

XX Query Match 91.8%; Score 928.6; DB 10; Length 2762;

XX Best Local Similarity 99.6%; Pred. NO. 4.5e-278;

XX Matches 931; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 77 AGATGGCTTCCCGAAGATGAGATGATATATCTGCTTCTGGTCTGGGGGCTCTTT 136

DB 432 AAATGGCTTCCCGAAGATGAGATGATATATCTGCTTCTGGTCTGGGGGCTCTTT 491

QY 137 GTTTGATTTTAGCATGTACAGTCTAAATCTCTTCAAGAACAGTCTTTTGTGTACAAGA 196

DB 492 GTTTGATTTTAGCATGTACAGTCTAAATCTCTTCAAGAACAGTCTTTTGTGTACAAGA 551

QY 197 AAGCGGGAACCTTCTTAAGCTTCCAGATACAGATGCGAGGAGACACCTCTCTCTCG 256

DB 552 AAGCGGGAACCTTCTTAAGCTTCCAGATACAGATGCGAGGAGACACCTCTCTCTCG 611

QY 257 TCCTGCTGGTGACCTCATCCACAAAACAGTTGGCTGAGCGCATGGCCATCCGGCAGAGCT 316
 DB 612 TCCTGCTGGTGACCTCATCCACAAAACAGTTGGCTGAGCGCATGGCCATCCGGCAGAGCT 671
 QY 317 GGGGAAAGAGGAGGAGCGGTGAAGGGAAGAGCTGAAGACATTTCTTCTCTGGGACCA 376
 DB 672 GGGGAAAGAGGAGGAGCGGTGAAGGGAAGAGCTGAAGACATTTCTTCTCTGGGACCA 731
 QY 377 CCAGCAGTGCAGCGGAAACAAAAGAGGTGGACAGGAGAGCCAGGACACCGGGACATTA 436
 DB 732 CCAGCAGTGCAGCGGAAACAAAAGAGGTGGACAGGAGAGCCAGGACACCGGGACATTA 791
 QY 437 TCCAGAGGATTTCTTAGACGTCTATTACATCTGACCTCAAGACCATGATGGCATAG 496
 DB 792 TCCAGAGGATTTCTTAGACGTCTATTACATCTGACCTCAAGACCATGATGGCATAG 851
 QY 497 AATGGGTCCATCGCTTTTGTCTCAGCGCGGTGTGTGATGATGAAAAACAGACTCAGACATGT 556
 DB 852 AATGGGTCCATCGCTTTTGTCTCAGCGCGGTGTGTGATGATGAAAAACAGACTCAGACATGT 911
 QY 557 TCATCAATGTTGACTATCTGACTGAATCTGCTTCTGAAGAAAAACAGAACCAACAGGTTTTT 616
 DB 912 TCATCAATGTTGACTATCTGACTGAATCTGCTTCTGAAGAAAAACAGAACCAACAGGTTTTT 971
 QY 617 TCATGGCTTCTTGAAACTCAATGAGTTTCCATCAGGCAGCCATTCAAGCAAGTGGTTTG 676
 DB 972 TCATGGCTTCTTGAAACTCAATGAGTTTCCATCAGGCAGCCATTCAAGCAAGTGGTTTG 1031
 QY 677 TCAGTAAATCTGAATATCCGTGGGACAGGTACCCACCATTTCTGCTCCGGCACCGGCTACG 736
 DB 1032 TCAGTAAATCTGAATATCCGTGGGACAGGTACCCACCATTTCTGCTCCGGCACCGGCTACG 1091
 QY 737 TGTTTTCTGGCGAGCTGCGAGTCAAGTGTACATGCTCCAAAGAGCTCCCATACATTA 796
 DB 1092 TGTTTTCTGGCGAGCTGCGAGTCAAGTGTACATGCTCCAAAGAGCTCCCATACATTA 1151
 QY 797 AACTGGAAGAGCTGTTGTGGGGCTCTGCCCTCGAAAGGCTGAACATCAGATTGGAGGAGC 856
 DB 1152 AACTGGAAGAGCTGTTGTGGGGCTCTGCCCTCGAAAGGCTGAACATCAGATTGGAGGAGC 1211
 QY 857 TCCACTCCAGCCGACCTTTTCCAGGGGCTTACGCTTCTCCGTATGCTCTTTTCAGGA 916
 DB 1212 TCCACTCCAGCCGACCTTTTCCAGGGGCTTACGCTTCTCCGTATGCTCTTTTCAGGA 1271
 QY 917 GGATCGTGGCTGCCACTTCAATCAAGCTTCGACTCTTTGGACTACTTGGCAGGCTCTAG 976
 DB 1272 GGATCGTGGCTGCCACTTCAATCAAGCTTCGACTCTTTGGACTACTTGGCAGGCTCTAG 1331
 QY 977 AGAATTCCTCCGGGGAAGATTGTCCGCTGTCTGA 1011
 DB 1332 AGAATTCCTCCGGGGAAGATTGTCCGCTGTCTGA 1366

RESULT 7

AAA93875

ID AAA93875 standard; DNA; 2775 BP.

XX AC AAA93875;

XX DT 15-JAN-2001 (first entry)

XX Human beta-1,3 galactose transferase encoding DNA.

XX Beta-1,3 galactose transferase; treatment; diagnosis; cancer; human;
 KW digestive system; ds.

XX Homo sapiens.

XX WO2000050608-A1.

XX PD 31-AUG-2000.

XX PF 24-FEB-2000; 2000WO-JP001070.

XX 25-FEB-1999; 99JP-00047571.
 XX (KYOW) KYOWA HAKKO KOGYO KK.
 XX Narimatsu H, Ishihiki S, Togayachi A, Sasaki K;
 XX WPI; 2000-549409/50.
 XX P-PSDB; AAB93875.
 XX Beta-1,3 galactose transferase and DNA encoding it, useful for synthesis
 XX of type 1 sialyl Lewis, a carbohydrate for treatment of digestive system
 XX cancer.
 XX Claim 5; Page 99-102; 123pp; Japanese.
 XX This invention relates to a polypeptide (I) with beta-1,3 galactose
 XX transferase activity, or variants of (I) comprising amino acid additions,
 XX deletions and/or substitutions. Included in the invention is DNA encoding
 XX all or part of (I); expression vectors containing the DNA, host cells
 XX transformed by the vectors; a method for the preparation of the
 XX polypeptide by culture of the transformants or by expression in the milk
 XX of a transgenic mammal, and antibodies recognising (I). The Beta-1,3
 XX galactose transferase protein transfers galactose by beta-1,3 bonding to
 XX N-acetylglucosamine present in a non-cyclic carbohydrate chain (such as
 XX GlcNAc beta1-3Gal beta1-4Glc) to give Gal beta1-3GlcNAc. The protein and DNA
 XX encoding it are useful for the treatment and diagnosis of cancer of the
 XX digestive system. The present sequence represents Beta-1,3 galactose
 XX transferase encoding DNA

Query Match 91.8%; Score 928.6; DB 3; Length 2775;
 Best Local Similarity 99.6%; Pred. No. 4.5e-278;
 Matches 931; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

77 AGATGGCTTTCCCGAAGATGAGATGATATATCTGCTTCTGCTTCTGGGGCTCTTT 136
 400 AAATGGCTTTCCCGAAGATGAGATGATATATTTGCTTCTGCTTCTGGGGCTCTTT 459
 137 GTTTGATTTTACGATGACAGTCTAAATCTTTTCAAGAACAGTCTTTGTTTACAGA 196
 460 GTTTGATTTTACGATGACAGTCTAAATCTTTTCAAGAACAGTCTTTGTTTACAGA 519
 197 AAGACGGGAACCTCTTAAAGTCCAGATACAGACTGACGAGCAGACCTCTCTTCG 256
 520 AAGACGGGAACCTCTTAAAGTCCAGATACAGACTGACGAGCAGACCTCTCTTCG 579
 257 TCCTGCTGGTGACCTCATCCCAACAAACAGTTGGCTGAGCGCATGCCATCCGCGAGCGT 316
 580 TCCTGCTGGTGACCTCATCCCAACAAACAGTTGGCTGAGCGCATGCCATCCGCGAGCGT 639
 317 GGGGGAAGAGAGAGCGTGAAGGAAGAGAGAGTCTTCTCTCTCTGGGACCA 376
 640 GGGGGAAGAGAGAGTGGTGAAGGGAAGAGAGTCTTCTCTCTCTGGGACCA 699
 377 CCAGCAGTGCAGCGAAGAAACAAAGAGTGGACAGGAGCAGCAGCAGCGGACATTA 436
 700 CCAGCAGTGCAGCGAAGAAACAAAGAGTGGAGAGGAGCAGGAGCAGCGGACATTA 759
 437 TCCAGAAGGATTTCTAGACGCTCTATTACAATCTGACCTGAAGACCATGATGGGCATAG 496
 760 TCCAGAAGGATTTCTAGACGCTCTATTACAATCTGACCTGAAGACCATGATGGGCATAG 819
 497 AATGGTCCATCGTTTGTCTCAGCGGGGCTTTGATGATGAAGAACAGACTCAGACATGT 556
 820 AATGGTCCATCGTTTGTCTCAGCGGGGCTTTGATGATGAAGAACAGACTCAGACATGT 879
 557 TCATCAATGTTGACTATCTGACTGAACTCTTCTGAGAAAACAGACACACAGGTTTT 616
 880 TCATCAATGTTGACTATCTGACTGAACTCTTCTGAGAAAACAGACACACAGGTTTT 939
 617 TCACTGGGCTTCTTGAAACTCAATGAGTTTCCCATCAGGCGAGCCATTTCAGCAAGTGTG 676

Db 940 TCACTGGGCTTCTTGAAACTCAATGAGTTTCCCATCAGGCGAGCCATTTCAGCAAGTGTG 999
 Qy 677 TCAGTAAATCTGAATATCCGTTGGGACAGGTACCCACCAATTCCTGCTCCGSCACCGCTAGG 736
 Db 1000 TCAGTAAATCTGAATATCCGTTGGGACAGGTACCCACCAATTCCTGCTCCGSCACCGCTAGG 1059
 Qy 737 TGTTTTCTGGCGACGCTGGCGAGTCAAGTGTACAATGTCTCCAAGAGCGTCCCATACATTA 796
 Db 1060 TGTTTTCTGGCGACGCTGGCGAGTCAAGTGTACAATGTCTCCAAGAGCGTCCCATACATTA 1119
 Qy 797 AACTGGAAGAGCTGTTGTTGGGGCTCTGCTTCGAAAGGCTGAAACATCAGATTGGAGGAGC 856
 Db 1120 AACTGGAAGAGCTGTTGTTGGGGCTCTGCTTCGAAAGGCTGAAACATCAGATTGGAGGAGC 1179
 Qy 857 TCCACTCCCGACGACCTTTTTCAGGGGCTTACGCTTCTCGTATGCTCTTTCAGGA 916
 Db 1180 TCCACTCCCGACGACCTTTTTCAGGGGCTTACGCTTCTCGTATGCTCTTTCAGGA 1239
 Qy 917 GGATCGTGGCTGCCACTTTCATCAAGGCTCGGACTCTTTCGAGCTACTGGCAGGCTCTAG 976
 Db 1240 GGATCGTGGCTGCCACTTTCATCAAGGCTCGGACTCTTTCGAGCTACTGGCAGGCTCTAG 1299
 Qy 977 AGAATTCGGGGGGAAGATTGTCGGCTGTCTGA 1011
 Db 1300 AGAATTCGGGGGGAAGATTGTCGGCTGTCTGA 1334

RESULT 8
 ADK68456
 ID ADK68456 standard; cDNA; 933 BP.
 XX AC ADK68456;
 XX DT 06-MAY-2004 (first entry)
 XX DE Human beta-1,3-galactosyl transferase cDNA SeqId 1.
 XX KW human; gene; ss: saccharide binding protein; maltose binding protein;
 XX KW beta-1,3-galactosyl transferase; inflammation; infectious disease;
 XX KW cancer metastasis suppression; dairy product; antiinflammatory;
 XX KW antimicrobial; cytostatic.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 XX CDS 1..933
 XX FT /*tag= a
 XX FT /product= "Beta-1,3-galactosyl transferase protein"
 XX PN JP2004016117-A.
 XX PD 22-JAN-2004.
 XX PF 17-JUN-2002; 2002JP-00176132.
 XX PR 17-JUN-2002; 2002JP-00176132.
 XX PA (TOYM) TOYOBO KK.
 XX WIPI; 2004-113878/12.
 XX P-PSDB; ADK68457.
 XX PT New fusion protein useful in diagnosis of diseases and in manufacture of
 XX PT pharmaceutical products, comprises a saccharide binding protein and beta
 XX PT 1, 3-galactosyl transferase.
 XX PS Example 1; SEQ ID NO 1; 23pp; Japanese.
 XX CC This invention relates to a novel recombinant fusion protein that
 XX CC comprises a saccharide binding protein, in particular a maltose binding
 XX CC protein, and a beta-1,3-galactosyl transferase and an appropriate
 XX CC manufacturing method. Specifically, it refers to fusion protein that can

CC transfer galactose to an N-acetyl glucosamine residue or N-acetyl
 CC glucosamine monosaccharide. The present invention describes a method to
 CC produce this enzyme, cheaply and efficiently, for the diagnosis and
 CC treatment of inflammation, infectious diseases or for cancer metastasis
 CC suppression. Furthermore, it can also be useful for improving dairy
 CC products. Accordingly, compositions exhibit various activities including
 CC antiinflammatory, antimicrobial and cytostatic. This polynucleotide
 CC sequence is the human beta-1,3-galactosyl transferase cDNA of the
 CC invention.

XX SQ Sequence 933 BP; 227 A; 232 C; 241 G; 233 T; 0 U; 0 Other;

Query Match		91.8%;	Score 928.2;	DB 12;	Length 933;
Best Local Similarity		99.7%;	Pred. No. 3.1e-278;		
Matches 930;		Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;
QY	79	ATGGCTTTCCCGAAGATGAGATTGATATATCTGCTTCTGGTCTCGGGGCTCTTTGT	138		
Db	1	ATGGCTTTCCCGAAGATGAGATTGATATATTTGCTTCTGGTCTCGGGGCTCTTTGT	60		
QY	139	TGTATTTTAGCATGTACAGTCTAAATCTTCAAGAACAGTCTTTGTTTACAGAAA	198		
Db	61	TGTATTTTAGCATGTACAGTCTAAATCTTCAAGAACAGTCTTTGTTTACAGAAA	120		
QY	199	GACGGAACTTCTTAAAGCTCCAGATACAGACTCGAGGACAGACCTCCCTTCTCTGTC	258		
Db	121	GACGGAACTTCTTAAAGCTCCAGATACAGACTCGAGGACAGACCTCCCTTCTCTGTC	180		
QY	259	CTGCTGGTGACCTCATCCCAAAACAGTTGGCTGAGCGCATCGGCATCCGGCAGACGTGG	318		
Db	181	CTGCTGGTGACCTCATCCCAAAACAGTTGGCTGAGCGCATCGGCATCCGGCAGACGTGG	240		
QY	319	GGGAAGAGAGAGAGCGTGAAGGAAAGCAGCTGAAGACATTTCTCTCTGGGACACACC	378		
Db	241	GGGAAGAGAGAGAGTGGTGAAGGAAAGCAGCTGAAGACATTTCTCTCTGGGACACACC	300		
QY	379	AGCAGTGCAGCGGAAACAAAGAGGTGACACAGGAGAGCGACCGGAGCAATATTC	438		
Db	301	AGCAGTGCAGCGGAAACAAAGAGGTGACACAGGAGAGCGACCGGAGCAATATTC	360		
QY	439	CAGAAGGATTTCTTAGAGCTTATTACAACTTGACCTTGAAAGACCATGATGGGCATAGAA	498		
Db	361	CAGAAGGATTTCTTAGAGCTTATTACAACTTGACCTTGAAAGACCATGATGGGCATAGAA	420		
QY	499	TGGGTCCATCGCTTTTGTCTCAGCGCGGCTTTGTGATGAAACAGACTCAGACATGTC	558		
Db	421	TGGGTCCATCGCTTTTGTCTCAGCGCGGCTTTGTGATGAAACAGACTCAGACATGTC	480		
QY	559	ATCAATGTTGACTATCTGACTGAATCTGCTTCTGAAGAAAAACAGAACCAACAGGTTTTC	618		
Db	481	ATCAATGTTGACTATCTGACTGAATCTGCTTCTGAAGAAAAACAGAACCAACAGGTTTTC	540		
QY	619	ACTGGCTTTCTTGAACACTCAATGAGTTTCCCATCAGGCGACCATTCAGCAAGTGGTTGTC	678		
Db	541	ACTGGCTTTCTTGAACACTCAATGAGTTTCCCATCAGGCGACCATTCAGCAAGTGGTTGTC	600		
QY	679	AGTAAATCTGAATATCCGTGGGACAGGTACCCACCATTTCTGCTCGGACACCGGCTACGTG	738		
Db	601	AGTAAATCTGAATATCCGTGGGACAGGTACCCACCATTTCTGCTCGGACACCGGCTACGTG	660		
QY	739	TTTTCTGCGGACGTGGCGAGTCAGGTGTACATGTCTCCAGAGCGTCCCATACATTAATA	798		
Db	661	TTTTCTGCGGACGTGGCGAGTCAGGTGTACATGTCTCCAGAGCGTCCCATACATTAATA	720		
QY	799	CTGGAAGACGTGTTTGTGGGCTCTGCTCCGCTCGAAAGGCTGAACATCAGATTGGAGGAGCTC	858		
Db	721	CTGGAAGACGTGTTTGTGGGCTCTGCTCCGCTCGAAAGGCTGAACATCAGATTGGAGGAGCTC	780		
QY	859	CACCTCCAGCGACCTTTTTCAGGGGGCTTACGCTTCTCGTATGCTCTTACGAGG	918		
Db	781	CACCTCCAGCGACCTTTTTCAGGGGGCTTACGCTTCTCGTATGCTCTTACGAGG	840		
QY	919	ATCGTGGGCTGCGCACTTTCATCAAGCCTCGGACTCTCTTTGGACTACTGCGAGGCTCTAGAG	978		

Db	841	ATCGTGGCTGCCACTTCATCAAGCTCGGACTCTCTTGGACTACTGGCAGGCTCTAGAG	900
QY	979	AATTCCTGGGGGGAAGATTGTCCGCTGTCTGA	1011
Db	901	AATTCCTGGGGGGAAGATTGTCCGCTGTCTGA	933
RESULT 9			
ACH79926	ID	ACH79926 standard; DNA; 596 BP.	
XX	AC	ACH79926;	
XX	DT	29-JUL-2004 (first entry)	
XX	DE	Human genome derived single exon probe #13121.	
XX	Human; probe; ss; gene expression; single exon probe; microarray;		
KW	alternative splicing event; genomic alteration.		
XX	Homo sapiens.		
OS			
XX	US2003194704-A1.		
PN	16-OCT-2003.		
XX			
XX	03-APR-2002; 2002US-00029386.		
PF			
XX	03-APR-2002; 2002US-00029386.		
PR			
XX	(PENN/) PENN S G.		
PA	(RANK/) RANK D R.		
PA	(HANZ/) HANZEL D K.		
XX			
PI	Penn SG, Rank DR, Hanzel DK;		
XX			
DR	WPI; 2004-119264/12.		
XX			
XX	New human genome-derived single exon nucleic acid probes useful for human		
PT	gene expression analysis, for identifying or characterizing alternative		
PT	splicing events, for assessing genomic alterations or as tools for		
PT	surveying tissues.		
XX			
PS	Claim 15; SEQ ID NO 13121; 80pp; English.		
XX			
CC	The invention relates to a nucleic acid probe for measuring human gene		
CC	expression, comprising any of the 27,400 fully defined nucleotide		
CC	sequences in the specification, or their complements or fragments, and		
CC	encoding at least 8 amino acids of any of the 6888 amino acid sequences		
CC	fully defined in the specification. The probe is a single exon probe that		
CC	hybridises under high stringency conditions to a nucleic acid molecule		
CC	expressed in human cells or tissues. Also included are a spatially-		
CC	addressable set of single exon nucleic acid probes for measuring human		
CC	gene expression (comprising a plurality of single exon nucleic acid		
CC	probes cited above, where each of the plurality of probes is separately		
CC	and addressably isolatable or amplifiable from the plurality), a single		
CC	exon microarray for measuring human gene expression, a method of		
CC	measuring human gene expression, a vector comprising the single exon		
CC	probe cited above, an ORF-encoded peptide comprising at least 8		
CC	contiguous amino acids of any of the above-mentioned amino acid		
CC	sequences (optionally with conservative amino acid substitutions), an		
CC	isolated antibody that binds specifically to a peptide cited above,		
CC	methods of selling and/or licensing single exon probes or microarrays to		
CC	a customer desiring to measure gene expression, a method of providing		
CC	human gene expression data by subscription, and a computer-readable		
CC	storage medium which contains a database having a plurality of records		
CC	(each record including data on the expression of a single exon probe		
CC	cited above. The probe, methods and apparatus are useful in gene		
CC	expression analysis. The probes may be used as tools for surveying		
CC	tissues to detect the presence of expressed messages that contain their		
CC	specific exon, or in constructing genome-derived single exon microarrays.		
CC	In addition, the probes are used in identifying and characterising		

CC alternative splicing events, in detecting and characterising gross
 CC alterations in the genomic locus that includes their exon, in assessing
 CC smaller genomic alterations, in priming the synthesis of nucleic acids,
 CC or in expressing the ORF-encoded peptide. The present sequence is a human
 CC single exon probe of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?DocID=20030194704
 XX
 SQ Sequence 596 BP; 149 A; 150 C; 141 G; 156 T; 0 U; 0 Other;
 Query Match 40.1%; Score 405.2; DB 12; Length 596;
 Best Local Similarity 99.3%; Pred. No. 3.1e-115;
 Matches 407; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 XX
 1 CCACCTCAGCTCCCTAGCATATAAACTAGACACATCTCTATGTTTGGTCTAATCATTT 60
 DB 187 CCACCTCAGCTCCCTAGCATATAAACTAGACACATCTCTATGTTTGGTCTAATCATTT 246
 QY 61 GGATTTTGTCTTTCAGATGCTTTCCGAAAGATGAGATTGATGTATATCTGCTTCTG 120
 DB 247 GGATTTTGTCTTTCAGATGCTTTCCGAAAGATGAGATTGATGTATATTTGCTTCTG 306
 QY 121 GTTCTGGGGCTCTTGTGTTTATTTAGCATGTACAGTCTAAATCTTTCAAAGACAG 180
 DB 307 GTTCTGGGGCTCTTGTGTTTATTTAGCATGTACAGTCTAAATCTTTCAAAGACAG 366
 QY 181 TCTTTTGTTTTACAAGAGACGGGAATCTTCTTAAGTCCCAAGATACAGACTGACGGCAG 240
 DB 367 TCTTTTGTTTTACAAGAGACGGGAATCTTCTTAAGTCCCAAGATACAGACTGACGGCAG 426
 QY 241 ACACCTCCCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
 DB 427 ACACCTCCCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 486
 QY 301 GCATCCCGCAGACGCTGGGGGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATTC 360
 DB 487 GCATCCCGCAGACGCTGGGGGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATTC 546
 QY 361 TTCTCTCTGGGACACACAGCAGTGCAGCGGAAACAAAAGAGGTTGACCA 410
 DB 547 TTCTCTCTGGGACACACAGCAGTGCAGCGGAAACAAAAGAGGTTGACCA 596

RESULT 10

ABZ56909

ID ABZ56909 standard; DNA; 1037 BP.

XX AC ABZ56909;

XX 04-APR-2003 (first entry)

XX DE B3GALT nucleic acid sequence # SEQ ID 1.

XX KW B3GALT; beta-1,3 galactosyltransferase; p53; cancer; breast; colon;

XX KW kidney; lung; ovary; gene; ds.

XX OS Homo sapiens.

XX FH Key

XX FT CDS

XX FT Location/Qualifiers

XX FT 35..1015

XX FT /*tag= a

XX FT /product= "B3GALT protein"

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX (EXEL-) EXELIXIS INC.
 XX Friedman L, Plozman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
 XX WPI; 2003-156849/15.
 XX P-PSDB; ABP60061.
 XX Identifying p53 pathway modulating agents with B3GALT genes, useful for
 XX the diagnosis and treatment of disorders associated with defects in the
 XX p53 pathway, such as cancer of the breast, colon, kidneys, lung and
 XX ovary.
 XX Example 3; Page 40; 82pp; English.
 XX The invention relates to identifying a candidate p53 pathway modulating
 XX agent in humans that is referred to in the specification as B3GALT (beta-
 XX 1,3 galactosyltransferase). Also disclosed is a method for diagnosing a
 XX disease in a patient, by contacting a sample with a probe for B3GALT
 XX expression, and comparing the results with a control, and determining
 XX whether the results indicate a likelihood of disease. Methods and
 XX compositions of the invention are useful for the diagnosis and treatment
 XX of disorders associated with defects in the p53 pathway, such as cancer
 XX of the breast, colon, kidneys, lung and ovary. The current sequence
 XX represents a B3GALT nucleic acid sequence referred to in an example from
 XX the invention
 XX SQ Sequence 1037 BP; 298 A; 240 C; 213 G; 286 T; 0 U; 0 Other;
 Query Match 16.2%; Score 163.4; DB 10; Length 1037;
 Best Local Similarity 55.7%; Pred. No. 1.1e-39;
 Matches 358; Conservative 0; Mismatches 276; Indels 9; Gaps 2;
 QY 206 ACTTCTTAAGCTCCAGATACAGACTGCGAGCAGACACCTCCCTCTCTGCTGCTGCTG 265
 DB 225 AATTTCCTTATCAACGAGCCCAATAATGTGAGAAAAACATTCCTTTCTTTGTTATCTCA 284
 QY 266 TGACCTCATCCCAACAAACAGTTGGCTGAGCGCATCGGCATCGGAGACGCTGGGGGAAAG 325
 DB 285 TCAGCACCACCTCACAAGGAATTTGATGCCGTGAGCAATCAGAGAGAGCTGGGGGATG 344
 QY 326 AGAGAGCGGTGAAGGAAAGCAGCTGAAGACATTTCTCTCTGGGGACCAACAGCAGGTG 385
 DB 345 AGAACAACTTTAAGGGGATCAAGATAGCCACCTGTTCTCTCTGGGCAAGATGCTGATC 404
 QY 386 CAGCGGAACAAAGAGGTGGACAGAGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 445
 DB 405 CTGTTCTCAATCAGATGTTGGAGCAAGAGAGCCAAATCTTCATGATATCATCTGTTGAGG 464
 QY 446 ATTTCTAGAGCTCTATTACATCTGACCTGAGACCATGATGGCATAGAAATGGTCC 505
 DB 465 ACTTATTGACTCTCTACCTAACCTTACCTTCAAAACATTAATGGGATGAGATGGTGG 524
 QY 506 ATCGCTTTTGTCTCAGCGGCGTTTGTGATGAAACAGACTCAGACATGTTTCATCAATG 565
 DB 525 CCACCTTTTGTCAAAGCCCAAGTATGTGTGAAACAGACAGAGCAGACATTTTGTAAACA 584
 QY 566 TTGACTATCTGACTGAATGCTTCTGAA-----GAAAAACAGAACACAGGTTTTC 619
 DB 585 TGGACAACTCTTATTATTAATAATTAATCTGAAACCCCTCCACCAAGCCAGAAAGGATTTT 644
 QY 620 CTGGCTTCTTGAACACTCAATGATTTTCCCATCAGCAGCCATTCAGCAAGTGGTTTGTCA 679
 DB 645 CTGGCTATGTCA---TTAATGAGACCGATTCGGGATGTCGGCAGTAAAGTATATGTC 701
 QY 680 GTAAATCTGAATATCCGTGGGACAGGTACCCACCATTTCTGCTCCGGCACCCTGCTGT 739
 DB 702 CCAGGGATTTGTACCCAGACAGTAACCTTACCCACCTTTCTGTTGGGGAGCTGCTACATCT 761
 QY 740 TTTCTGGGAGCTGGCGAGTCAAGTGTACATGTTCTCCAGAGCGTCCATACATTAAC 799
 DB 762 TTTGAGCGAGTGTAGCTGAATCTATTTTACAAGACCTCTACTCCACACAAGGCTGCTTCC 821

QY 800 TGAAGACGCTGTTTGGGGCTCTGCTCGAAAGGCTGAACAT 842
 |||||
 DB 822 TTGAAGACGTATATGTGGGACTGTGCTTCGAAGCTGGGCAT 864
 |||||

RESULT 11
 AAQ67067
 ID AAQ67067 standard; cDNA to mRNA; 1739 BP.
 XX
 AC AAQ67067;
 XX
 DT 14-MAR-1995 (first entry)
 XX
 DE Beta-1,3-galactosyltransferase cDNA.
 XX
 KW Beta-1,3-galactosyltransferase; enzyme; saccharide chain; pMoPRW1;
 KW KJM-1 cells; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 676..1656
 FT /*tag= a
 FT /product= "Beta-1,3-galactosyltransferase"
 XX

PN JP06181759-A.
 XX
 PD 05-JUL-1994.
 XX
 PF 16-DEC-1992; 92JP-00336436.
 XX
 PR 16-DEC-1992; 92JP-00336436.
 XX
 PA (KYOW) KYOWA HAKKO KOGYO KK.
 XX
 DR WPI; 1994-251683/31.
 XX
 DR P-PSDB; AAR57433.
 XX
 XX Beta-galactosyl-transferase DNA and protein - useful for prodn. of
 FT saccharide chains.
 XX
 PS Claim 3; Page 22-24; 47pp; Japanese.
 XX
 CC This sequence encodes a beta-1,3-galactosyltransferase. This enzyme can
 CC be used to produce physiologically active saccharide chains and variants,
 CC and for improvement of saccharide chains bound to physiologically active
 CC proteins. This cDNA represents a fragment of the plasmid pMoPRW1 which
 CC was cloned in KJM-1 cells
 XX
 SQ Sequence 1739 BP; 473 A; 407 C; 422 G; 437 T; 0 U; 0 Other;

Query Match 16.2%; Score 163.4; DB 2; Length 1739;
 Best Local Similarity 55.7%; Pred. No. 1.5e-39;
 Matches 358; Conservative 0; Mismatches 276; Indels 9; Gaps 2;
 QY 206 ACTTCCTTAAGCTCCAGATACAGACTCGAGGACACCTCCCTCTCTGCTGCTGG 265
 |||||
 DB 866 AATTCTTATCAACGAGCCCAATAAATGTGAGAAAAACATTCCTTTCTGTATCTCA 925
 |||||
 QY 266 TGACCTCATCCACAAAGTGGCTGAGCGCATGGCCATCCGGCAGACGTGGGGGAAAG 325
 |||||
 DB 926 TCAGACCACTCACAAGGAATTTGATGCCGTGAGCAATCAGAGACGTGGGGGATG 985
 |||||
 QY 326 AGAGACCGGTGAAGGAAAGAGCTGAAGACATTTCTCTCTGGGACCAACACAGAGTG 385
 |||||
 DB 986 AGAACAACTTTAAGGGGATCAAGATAGCACCTCTCTCTGGGCAAGAATGCTATC 1045
 |||||
 QY 386 CAGCGGAACAAAGAGCTGGACGAGAGCCAGCAGACATTCATCCAGAGG 445
 |||||
 DB 1046 CTGTCTCAATCAGATGGTGGAGCAAGAGCCAAATCTTCATGATATCATCGTGGAGG 1105
 |||||
 QY 446 ATTTCTAGACGTCTATTACATCTGACCTGAAGACCATGATGGCATAGATGGGTCC 505
 |||||

DB 1106 ACTTTATGACTCTCTACCAATAACCTTTACCTCAAAACATTAAATGGGATGAGATGGGTG 1165
 |||||
 QY 506 ATCGCTTTTGTCTCTCAGGGCGGCTTTTGTGATGAAAAACAGACTCAGACATGTTTCATCAATG 565
 |||||
 DB 1166 CCACTTTTGTTCAAAAGCCAAGTATGTCATGAAAACAGACAGCGACATTTTGTAAACA 1225
 |||||
 QY 566 TTGACTATCTGACTGAACCTGCTTCTGAA-----GAAAAACAGAAACAACAGGTTTTTCA 619
 |||||
 DB 1226 TGGACAATCTTATTATTAATAATTACTGAAACCTCCACCAAGCCACGAAGAAGGATTTTA 1285
 |||||
 QY 620 CTGGCTTCTTGAAGAACTCAATGAGTTTCCCATCAGGAGCCATTCAGCAAGTGGTTGTCA 679
 |||||
 DB 1286 CTGGCTATGTCA---TTAATGGAGGACCGATTCGGGATGTCGCGAGTAAGTGTATATGC 1342
 |||||
 QY 680 GTAAATCTGAATATCCGTGGGACAGGTACCCACATTTCTGCTCCGGCACCGGCTACGTGT 739
 |||||
 DB 1343 CCAGGGAATTTGTATCCAGACAGTAATCACTTTTACAAGACCTCACTCCACACAAGGCTGCTTCA 1402
 |||||
 QY 740 TTCTTGGCGACGTGGCGAGTCAAGTGTACAATGTTCTCCAAGAGCGTCCCATATCAATTAAC 799
 |||||
 DB 1403 TTTTCAGCGGATGTAGCTGAACATTTTACAAGACCTCACTCCACACAAGGCTGCTTCA 1462
 |||||
 QY 800 TGAAGACGCTGTTTGGGGCTCTGCTCGAAAGGCTGAACAT 842
 |||||
 DB 1463 TTGAAGACGTATATGTGGGACTGTGCTTTCGAAAGCTGGGCAT 1505
 |||||

RESULT 12
 ABZ56910
 ID ABZ56910 standard; DNA; 2168 BP.
 XX
 AC ABZ56910;
 XX
 DT 04-APR-2003 (first entry)
 XX
 DE B3GALT nucleic acid sequence # SEQ ID 2.
 XX
 DE B3GALT; beta-1,3 galactosyltransferase; p53; cancer; breast; colon;
 KW kidney; lung; ovary; gene; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 352..1332
 FT /*tag= a
 FT /product= "B3GALT protein"
 XX
 XX WO200299044-A2.
 XX
 PD 12-DEC-2002.
 XX
 PF 02-JUN-2002; 2002WO-US017356.
 XX
 PR 05-JUN-2001; 2001US-0296076P.
 XX
 PR 10-OCT-2001; 2001US-0328605P.
 XX
 PR 15-FEB-2002; 2002US-0357253P.
 XX
 FA (EXEL-) EXELIXIS INC.
 XX
 PI Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
 XX
 DR WPI; 2003-156849/15.
 XX
 DR P-PSDB; ABP60062.
 XX
 PT Identifying p53 pathway modulating agents with B3GALT genes, useful for
 PT the diagnosis and treatment of disorders associated with defects in the
 PT p53 pathway, such as cancer of the breast, colon, kidneys, lung and
 PT ovary.
 XX
 PS Disclosure; Page 41-42; 82pp; English.
 XX
 CC The invention relates to identifying a candidate p53 pathway modulating
 CC agent in humans that is referred to in the specification as B3GALT (beta-

368	Qy	TGGGGACCA	CGCAGCTG	CGCGGAACA	AAAGAGG-----	-----TGG	ACCAGGAGCC	418				
449	Db	TAGGCCA	AGAGGCTG	AAAAAGGA	GACAAAAATG	TGGCATTTG	CTCTTAGAGGAT	GAACACC	508			
419	Qy	AGCGACAC	GGGGACAT	TATCCAGAA	GGATTTCTC	TAGACGTC	TATTACAA	TCTGCACCC	TGA	478		
509	Db	TTCTTTAT	TGGTGACAT	AAACCGAC	AGATTTTTT	TAGACACAT	ATATATAC	CTTACCTTGA	568			
479	Qy	AGACCAT	TGATGGGC	ATAGAAAT	TGGTCCAT	TGCTTTT	TGCTCAGG	CGGGCTTT	TGTGATGA	538		
569	Db	AAACCAT	TATGGCAT	TCAGGTGG	GTAACCTG	AGTTTGG	CCCCAATG	CCAAGTAC	GTAATGA	628		
539	Qy	AAACAGAC	TCTCAGAC	ATGTTTCA	TCAATGTTG	ACTATCTG	ACTGACTG	CACTGCTT	CTCAAGNAAA	598		
629	Db	AGACGACA	CACTGATG	TTTTCATC	AACTACTG	GGCAATTT	TAGTGA	AGATATCT	TTTAAACCTAA	688		
599	Qy	ACAGAA	CAACCCAG	GTTTTTT	TTCACCTG	GGCTTCTT	GTGAAACT	CAATGAGT	TTTCCCATCAGG	CAGC	658	
689	Db	ACCAC	TACAGAGA	AGTTTTT	TTCACAGG	TATCCTCT	TAATTG	ATATAT	TCTATAGAG	GAT	748	
659	Qy	CATTAC	GACAGT	GGTTT	TGTCAGT	TAATCTG	AAATCTG	AAATCCG	TGGGACAG	GTACCCACCA	TTCT	718
749	Db	TTTACCA	AAAAAAC	CCCATAT	TTCTTAC	CAGAGTAT	CCTTTC	CAAGGTG	TTCCCTCC	CATACT	808	
719	Qy	GCCTCCG	CCACCG	CTACGTG	TTTTTCTG	CGGACGTG	GGGAGT	CAAGGTG	TACAA	TGTCTCCA	778	
809	Db	GCAGT	GGTGGT	GTATATA	TAATGTC	CAGAGAT	TTTGGT	GCCAGG	ATCTATG	AAATGATG	868	
779	Qy	AGAGCG	TCCCAT	ACATTA	AAACTG	GAAGAC	GTGTTG	TGTGGG	CTCTCG	CTCGAAAG	CGCTGA	838
869	Db	GTACG	GTAAAA	CCCATCA	AGTTTGA	AGATGTTAT	GTCTGG	GATCTG	TTTGA	ATTTATTA	928	
839	Qy	ACATCAG	ATGAGG	AGCTCCA	TCCACG	CCGACCT	TTTTTTT	TCCAGG	GGGGCTT	ACGCTTCT	898	
929	Db	AAGTGA	CAATTC	ATATAT	TCCAGA	GACACAAA	CTTTTCT	TCTTAT	TATAGA	ATCCATTTGG	988	
899	Qy	CCGTATG	CTCTCT	TCAGGAG	ATCGTG	CGCTGCC	ACTTCAT	CAAGCCT	TCGACT	CTCTTTGG	958	
989	Db	ATGCTG	TGTC	AAC	TGACG	ACGTGTG	ATTTG	CAGCC	CAATGGCT	TTTCTTCA	AGGAGATCATCA	1048
959	Qy	ACTACT	GGCAGG	970								
1049	Db	CTTTT	GGCAGG	1060								

RESULT 15

ADP76806

ID ADP76806 standard; cDNA; 1358 BP.

XX ADP76806;

AC

XX

XX

XX 26-FEB-2004 (first entry)

XX

XX Novel human secreted and transmembrane protein cDNA SeqID 481.

XX human; PRO; membrane bound protein; membrane bound receptor;

KW cell proliferation; cell migration; cell differentiation;

KW mitogenic factor; survival factor; cytotoxic factor;

KW differentiation factor; neuropeptide; hormone; cell receptor;

KW receptor-ligand interaction; cytosolic; chondrocyte; tumour; ss; gene.

XX

XX Homo sapiens.

XX

XX WO2003072035-A2.

XX

XX 04-SEP-2003.

XX

XX 21-FEB-2003; 2003WO-US005241.

XX

XX 22-FEB-2002; 2002US-0359461P.

XX

XX (GETH) GENENTECH INC.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2005, 12:15:17 ; Search time 210.477 Seconds
(without alignments)
7859.644 Million cell updates/sec

Title: US-10-777-828-8

Perfect score: 1011

Sequence: 1 ccactcagcctcctagcat.....aagatttcgcgtgtctga 1011

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_NA.*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PTCUS_COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1011	100.0	1011	4	US-09-831-630-8
2	114.8	11.4	1134	3	US-09-482-180A-3
3	78.4	7.8	1420	3	US-09-482-180A-1
4	78.4	7.8	22374	4	US-09-949-016-14938
5	70.2	6.9	1191	3	US-09-459-133-3
6	68.6	6.8	1167	3	US-09-459-133-14
7	67.4	6.7	1464	4	US-09-774-528-430
8	52.8	5.2	1434	2	US-09-055-097-2
9	52.8	5.2	1434	4	US-09-373-902-2
10	52.8	5.2	1640	4	US-09-949-016-5696
11	52.8	5.2	5640	4	US-09-949-016-17438
12	52.8	5.2	8436	4	US-09-949-016-13963
13	46	4.5	12902	4	US-09-949-016-13715
14	44.8	4.4	1532	3	US-09-459-133-1
15	43.4	4.3	1446	3	US-09-459-133-12
16	43.2	4.3	31166	4	US-09-949-016-13464
17	41.4	4.1	7218	1	US-08-233-463-14
18	39.2	3.9	748	4	US-09-270-767-444
19	39.2	3.9	748	4	US-09-270-767-15726
20	39.2	3.9	4079	4	US-09-016-434-1248
21	39.2	3.9	4517	4	US-09-949-016-4573
22	39.2	3.9	4519	4	US-09-023-655-1202
23	39.2	3.9	4519	4	US-09-949-016-650
24	39.2	3.9	4519	4	US-09-492-027-3
25	39.2	3.9	317366	4	US-09-949-016-16001
26	38.6	3.8	17731	4	US-09-949-016-16365
27	38.2	3.8	601	4	US-09-949-016-104020

ALIGNMENTS

RESULT 1

US-09-831-630-8
; Sequence 8, Application US/09831630
; Patent No. 6800468
; GENERAL INFORMATION:
; APPLICANT: Clausen, Henrik
; APPLICANT: Amado, Margarita
; TITLE OF INVENTION: UDP-GALACTOSE: BETA-N-ACETYL-GLUCOSAMINE BETA 1,3
; FILE REFERENCE: 7188-157
; CURRENT APPLICATION NUMBER: US/09/831.630
; CURRENT FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 8
; LENGTH: 1011
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (79) ..(1008)
US-09-831-630-8

Query Match	100.0%	Score	1011;	DB	4;	Length	1011;
Best Local Similarity	100.0%	Pred. No.	7.5e-313;	Mismatches	0;	Indels	0;
Matches	1011;	Conservative	0;	0;	0;	Gaps	0;
Qy	1	CCACCTCAGCCTCCTAGCATAAACTAGACACATCCTCATCTTTTGGAGGCTTAATCATTT	60				
Db	1	CCACCTCAGCCTCCTAGCATAAACTAGACACATCCTCATCTTTTGGAGGCTTAATCATTT	60				
Qy	61	GGATTTTGTTCCTTTCAGATGGCTTTCCGGAAGATGAGATTGATGATATATCTGCTTCG	120				
Db	61	GGATTTTGTTCCTTTCAGATGGCTTTCCGGAAGATGAGATTGATGATATATCTGCTTCG	120				
Qy	121	GTTCCTGGGGGCTCTTTGTTGTTATTTTAGCATGATAGTCTAAATCCTTTCAAGAAGCAG	180				
Db	121	GTTCCTGGGGGCTCTTTGTTGTTATTTTAGCATGATAGTCTAAATCCTTTCAAGAAGCAG	180				
Qy	181	TCCTTTGTTTACAGAAGACGGCACTTCCTTAAGCTCCAGATACAGATCTGAGGCGAG	240				
Db	181	TCCTTTGTTTACAGAAGACGGCACTTCCTTAAGCTCCAGATACAGATCTGAGGCGAG	240				
Qy	241	ACACCTCCCTTCCTCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	300				
Db	241	ACACCTCCCTTCCTCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	300				
Qy	301	GCCATCCGGCAGACGTGGGGGAAAGAGAGGACGGTGAAGGAAAGAGAGGAGGAGGAGG	360				
Db	301	GCCATCCGGCAGACGTGGGGGAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	360				

```

Qy 361 TTCTCTCTGGGACCAACAGCAGTGCAGCGAAGAAAGAGGTGGACAGAGCCAG 420
Db 361 TTCTCTCTGGGACCAACAGCAGTGCAGCGAAGAAAGAGGTGGACAGAGCCAG 420
Qy 421 CGACAGGGGACATTATCCAGAAGATTCTCTAGACGCTATTAACTCTGACCTGAAG 480
Db 421 CGACAGGGGACATTATCCAGAGAGATTCTTAGACGCTATTAACTCTGACCTGAAG 480
Qy 481 ACCATGATGGGCATAGATGGGTCCATCGCTTTTGTCTCAGCGCGGTGTGTGATGAAA 540
Db 481 ACCATGATGGGCATAGATGGGTCCATCGCTTTTGTCTCAGCGCGGTGTGTGATGAAA 540
Qy 541 ACAGACTCAGACATGTTTCATCAATGTTGACTACTCTGACTGAACTGCTCTGAGAAAAAC 600
Db 541 ACAGACTCAGACATGTTTCATCAATGTTGACTACTCTGACTGAACTGCTCTGAGAAAAAC 600
Qy 601 AGAACCAACAGGTTTTTTCACCTGGCTTCTTGAAACTCAATGAGTTTCCCATCAGCAGCCA 660
Db 601 AGAACCAACAGGTTTTTTCACCTGGCTTCTTGAAACTCAATGAGTTTCCCATCAGCAGCCA 660
Qy 720 TTACAGCAAGTGGTTGTTCAGTAAATCTGAAATATCCGTGGGACAGTACCCACCATTTCTGC 720
Db 720 TTACAGCAAGTGGTTGTTCAGTAAATCTGAAATATCCGTGGGACAGTACCCACCATTTCTGC 720
Qy 780 TCCGGCACCGGCTAGCTGTTTCTGGCGACGTCGCGAGTCAAGTGTACAATGTCTCCAG 780
Db 780 TCCGGCACCGGCTAGCTGTTTCTGGCGACGTCGCGAGTCAAGTGTACAATGTCTCCAG 780
Qy 840 AGCGTCCCATACATTAAACTGGAAGCTGTTTGTGGGGCTCTGCTCGAAGGCTGAAC 840
Db 840 AGCGTCCCATACATTAAACTGGAAGCTGTTTGTGGGGCTCTGCTCGAAGGCTGAAC 840
Qy 900 ATCAGATGGAGGAGCTCCACTCCAGCGACCTTTTTCAGGGGGCTTACGCTTCTCC 900
Db 900 ATCAGATGGAGGAGCTCCACTCCAGCGACCTTTTTCAGGGGGCTTACGCTTCTCC 900
Qy 960 GTATGCTCTTCAGGAGATCGTGCCCTGCCACTTCATCAAGCCCTCGGACTCTCTGGAC 960
Db 960 GTATGCTCTTCAGGAGATCGTGCCCTGCCACTTCATCAAGCCCTCGGACTCTCTGGAC 960
Qy 1011 TACTGGCAGGCTTAGAGAAATCCCGGGGGAAGATTGTCGCGCTGTCTGA 1011
Db 961 TACTGGCAGGCTTAGAGAAATCCCGGGGGAAGATTGTCGCGCTGTCTGA 1011

```

```

RESULT 2
US-09-482-180A-3
; Sequence 3, Application US/09482180A
; Patent No. 6361985
; GENERAL INFORMATION:
; APPLICANT: Konklin, Darrell C.
; APPLICANT: Yamamoto, Gayle
; APPLICANT: Gao, Zeren
; APPLICANT: Whitmore, Theodore E.
; APPLICANT: Jasper, Stephen
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING
; FILE REFERENCE: 98-80
; CURRENT APPLICATION NUMBER: US/09/482,180A
; PRIORITY FILING DATE: 2000-01-12
; PRIOR FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 1134
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: degenerate sequence
; NAME/KEY: variation
; LOCATION: (1)...(1134)

```

```

; OTHER INFORMATION: n is any nucleotide
; NAME/KEY: misc_feature
; LOCATION: (1)...(1134)
; OTHER INFORMATION: n = A,T,C or G
US-09-482-180A-3

```

Query Match 11.4%; Score 114.8; DB 3; Length 1134;

Best Local Similarity 31.3%; Pred. No. 7.1e-26; Mismatches 358; Indels 3; Gaps 1;

```

Matches 209; Conservative 98; Mismatches 358; Indels 3; Gaps 1;

Qy 181 TCCTTTCTTTTACAGAAAGACGGGAACCTCTTAAAGCTCCAGATACAGATCGAGCGAG 240
Db 286 TTYTNACNTAYMGNCAATGYMGNAAVTTYWSNATHYNTYNGARCCNWSNGNTGYWSN 345
Qy 241 ACACCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
Db 346 AARGAYACNTTYTNTYNTYNTYNTYNTYNTYNTYNTYNTYNTYNTYNTYNTYNTYNTY 405
Qy 301 GCCATCCGCGCAGAGCTGGGGGA--AAGAGAGGACGGTGAAGGAAAGAGCTGAAGACA 357
Db 406 GCNATHMGNWSNACNTGGGNGMNGNTGNGNGNTGNGNGMNGMNGNCARYTNAATYN 465
Qy 358 TTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 417
Db 466 GTNTTYTNTYNTYNTYNTYNTYNTYNTYNTYNTYNTYNTYNTYNTYNTYNTYNTYNT 525
Qy 418 CAGCGACACGCGGACATTTATCCAGAGGATTTCTCTAGACGCTTATTAATCTGACCCCTG 477
Db 526 MNGARTTYGAYGAYATHYTNCAATGCGGAYTTVACNGARGAYTTTYYAAAYTNAACNTN 585
Qy 478 AAGACCATGATGGGCATAGAAATGGGTCCATCGCTTTTGTCTCTCAGCGCGGCTTGTGATG 537
Db 586 AARGARYTNCAYTNCARMGNTGGGTNGTNGCNGTGYCCNCAAGCNCARYTATYATYTN 645
Qy 538 AAAACAGACTCAGACATGTTTCATCAATGTTTGAATCTCTGACTGAACTGCTTCTGAAGAAA 597
Db 646 AARGENGAYGAYGAYGAYGAYGAYGAYGAYGAYGAYGAYGAYGAYGAYGAYGAYGAY 705
Qy 598 AACAGAAACAACGAGTTTTCATCTGGCTTCTTGAACCTCAATGAGTTTCCATCAGCGAG 657
Db 706 GAYCCNCGNCARGAYTNTYNTYNTYNTYNTYNTYNTYNTYNTYNTYNTYNTYNTYNTY 765
Qy 658 CCATTACAGCAAGTGGTTTGTCAATCTGATAATCTCCGTGGGACAGGTACCCACCATTC 717
Db 766 ACNARTNARTAYTTTATTCNCCNWSNATGTAYMNGNCNACNATYNTYNTYNTYNTYNTY 825
Qy 718 TGCTCCGCGCACCGCTACGTTTCTTCTGGCGAGCTGGCGAGTCAAGTGTACAATGTCTCC 777
Db 826 CGNGGNGGNGGNTAYGTATGWSNMGNACNCTNNGMNGNYTNCARGCNCATATG 885
Qy 778 AAGAGCTCCCATACATTAACCTGGAAGACGTTTGTGGGGCTCTCGCTCGAAGAGCTG 837
Db 886 GARGAYCNGARYTNTYNTYNTYNTYNTYNTYNTYNTYNTYNTYNTYNTYNTYNTYNTY 945
Qy 838 AACATCAG 845
Db 946 GGNNTWS 953

```

```

RESULT 3
US-09-482-180A-1
; Sequence 1, Application US/09482180A
; Patent No. 6361985
; GENERAL INFORMATION:
; APPLICANT: Konklin, Darrell C.
; APPLICANT: Yamamoto, Gayle
; APPLICANT: Gao, Zeren
; APPLICANT: Whitmore, Theodore E.
; APPLICANT: Jasper, Stephen
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING
; FILE REFERENCE: 98-80
; CURRENT APPLICATION NUMBER: US/09/482,180A

```

```

; CURRENT FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/115,721
; FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1420
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (135)...(1271)
US-09-482-180A-1

Query Match
Best Local Similarity 7.8%; Score 78.4; DB 3; Length 1420;
Matches 299; Conservative 0; Mismatches 346; Indels 3; Gaps 1;

Qy 201 CGGGAATTCCTTAAGCTCCAGATACAGACTGCAGGCGAGACACCTCCCTTCCTGCTCCT 260
Db 440 CCGAAATTTCTCTATCTTGTCTGGAGCTTTCAGGCTGTTCAGGATACCTTCTTGTCTCT 499

Qy 261 GCTGTGACCTCATCCCAACAAAGTTGGCTGAGCGCATGCCATCCGGCAGAGCTGGG 320
Db 500 GGCCATCAAGTCACAGCTGTGTCAAGTGGAGCGTGGGCTATCCGACGACGCTGGG 559

Qy 321 GAAAGAGAGG---ACGGTGAAGGGAAGCAGCTGAAGACATTTCTTCTCTCTGGGGACAC 377
Db 560 CAGGTGGGGGATGGCTAGGGCCGCGAGCTGAAGCTGTGTCTCTCTAGGGGTGC 619

Qy 378 CAGCAGTCCAGGGGAAACAAAGAGTGGACAGGAGCCAGCGACACGCGGACATTA 437
Db 620 AGGATCCCTCCCGCAGCCAGCTCTGGCTATGAGAGTAGGGAGTTTGTATGACATCT 679

Qy 438 CCAGAGGATTTCTTAGAGCTTATTAACAATCTGACCTGAAGACCATATGATGGGCATAGA 497
Db 680 CCAGTGGGACTTCACTGAGGACTTCTTCAACCTGACGCTCAAGGAGCTGCACCTGCGAGCG 739

Qy 498 ATGGGTCCATCGCTTTTGTCTCAGCGCGCGTTTGTGATGAAACACAGACTCAGACATGTT 557
Db 740 CTGGGTGTGGCTGCTGCTGCCCCAGGCCATTTCTATGCTAAAGGAGATGACATGCTT 799

Qy 558 CATCAATGTTGACTATCTGACTGACTGCTTCTTGAAGAAACACAGAAACACAGGTTTT 617
Db 800 TGTCCACGTCGCCAACGTTAGAGTTCTGTGATGGCTGGGACCCAGCCAGGACCTCT 859

Qy 618 CACTGGCTTCTTGAACCTCAATGATTTCCCATCAGGCGAGCCATTCAGCAAGTGGTTGT 677
Db 860 GGTGGGAGATGTCACTCCGCCAAGCCCTGCCCAACAGGAACTAAGGTCAAAATACTTCAT 919

Qy 678 CAGTAAATCTGAATATCCGTGGGACAGGTACCCACATTTCTGCTCCGGCACCGGCTACGT 737
Db 920 CCCACCTCAATGTACAGGGCCACCACTACCCACCTATGCTGTGGGGAGGATATGT 979

Qy 738 GTTTTCTGGCGAGTGGCGAGTCAAGTGTAACAATGTCTTCCAGAGCGTCCCATACATPAA 797
Db 980 CATGTCCAGAGCCAAGTGCAGCGCGCTCCAGCTATCATGGAAGATGTGAATCTCTC 1039

Qy 798 ACTGGAACAGCTTTTGTGGGCTCTGCTCGAAGGCTGAACATCAG 845
Db 1040 CATTGATGATGTCTTTGTGGGTATGTGCTCAGGAGGCTGGGGCTGAG 1087

RESULT 4
US-09-949-016-14938/C
; Sequence 14938, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016

```

```

; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14938
; LENGTH: 22374
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(22374)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14938

Query Match
Best Local Similarity 7.8%; Score 78.4; DB 4; Length 22374;
Matches 299; Conservative 0; Mismatches 346; Indels 3; Gaps 1;

Qy 201 CGGGAATTCCTTAAGCTCCAGATACAGACTGCAGGCGAGACACCTCCCTTCCTGCTCCT 260
Db 21586 CCGAAATTTCTCTATCTTGTCTGGAGCCTTCAGGCTGTTCAGGATACCTTCTTGTCTCT 21527

Qy 261 GCTGTGACCTCATCCCAACAAAGTTGGCTGAGCGCATGCCATCCGGCAGAGCTGGG 320
Db 21526 GGCCATCAAGTCACAGCTGTGTCAAGTGGAGCGCTATCCGACGACGCTGGG 21467

Qy 321 GAAAGAGAGG---ACGGTGAAGGGAAGCAGCTGAAGACATTTCTTCTCTCTGGGGACAC 377
Db 21466 CAGGTGGGGGATGGCTAGGGCCGCGAGCTGAAGCTGTGTCTCTCTAGGGGTGC 21407

Qy 378 CAGCAGTGCAGCGGGAACAAAGAGTGGACCGAGAGCCAGCGACACGCGGACATTA 437
Db 21406 AGGATCCGCTCCCGCAGCCAGCTGTGGCTATGAGAGTAGGGAGTTTGTATGACATCT 21347

Qy 438 CCAGAGGATTTCTTAGAGCTTATTAACAATCTGACCTGAAGACCATATGATGGGCATAGA 497
Db 21346 CCAGTGGAGCTTCACTGAGGACTTCTTCAACCTCAGCGCTCAAGGAGCTGCACCTGCAGCG 21287

Qy 498 ATGGGTCCATCGCTTTTGTCTCAGCGCGCGTTTGTGATGAAACACAGACTCAGACATGTT 557
Db 21286 CTGGGTGTGGCTGCTGCTGCCCCAGGCCATTTCTATGCTAAAGGAGATGACATGCTT 21227

Qy 558 CATCAATGTTGACTATCTGACTGACTGCTTCTTGAAGAAACACAGAAACACAGGTTTT 617
Db 21226 TGTCCAGCTCCCAACGTTTAGAGTTCTCTGGATGGCTGGGACCCAGCCAGGACCTCT 21167

Qy 618 CACTGGCTTCTTGAACCTCAATGATTTCCCATCAGGCGAGCATTCAGCAAGTGGTTGT 677
Db 21166 GGTGGGAGATGTCTATCCCGAGCCCTGCCCAAGGAACTAAGGTCAAAATACTTCAT 21107

Qy 678 CAGTAAATCTGAATATCCGTGGGACAGGTACCCACCATTTCTGCTCCGGCACCCGCTACGT 737
Db 21106 CCCACCTCAATGTACAGGGCCACCACTACCCACCTATGCTGTGGGGAGGATATGT 21047

Qy 738 GTTTTCTGGCGAGCTGGCGAGTCAAGTGTAACAATGTCTTCCAGAGCGTCCCATACATPAA 797
Db 21046 CATGTCCAGAGCCACAGTGGCGCCCTCCAGGCTATCATGGAAGATGTGAATCTTCTCC 20987

Qy 798 ACTGGAAGAGCTGTGTGTGGGCTCTGCTCGAAGGCTGAACATCAG 845
Db 20986 CATTGATGATGTCTTTGTGGGTATGTGCTCAGGAGGCTGGGGCTGAG 20939

RESULT 5
US-09-459-133-3
; Sequence 3, Application US/09459133
; Patent No. 6416988
; GENERAL INFORMATION:

```


[illegible]

Db 9845 CGTCTCAACAAACAACAAACACCTTATATGGCATTCAGGTGGGTAACTAGATTG 9786
Qy 516 TCCTCAGGCGGGTGTGTGATGAACAGACTAGACATGTTTCATCAATGTGACTATCT 575
Db 9785 CCCATTGCCAATACATCATCATGACACAGACACTTATGTGTTTCATCAATACTGGCAATT 9726
Qy 576 GACTGAACCTGCTTCTGAAGAAACAGAACACAGTGTTCCTCACTGCTTCTTGAAC 635
Db 9725 AGTGAGTATCTTTTAAACTTAAACCACTCAGAGAAGTTTTTTCACAGGTTATCCCTTAAT 9666
Qy 636 CAATGAGTTTCCCATCAG 653
Db 9665 TGATAATATTCCCATAG 9648

RESULT 14
US-09-459-133-1
; Sequence 1, Application US/09459133
; Patent No. 6416988
; GENERAL INFORMATION:
; APPLICANT: Konklin, Darrell C.
; APPLICANT: Yamamoto, Gayle
; APPLICANT: Jaspers, Stephen R.
; APPLICANT: Gao, Zeren
; TITLE OF INVENTION: BETA-1,3-GALACTOSYLTRANSFERASE HOMOLOGS
; FILE REFERENCE: 98-77
; CURRENT APPLICATION NUMBER: US/09/459,133
; CURRENT FILING DATE: 1999-12-10
; PRIOR APPLICATION NUMBER: 60/111,697
; PRIOR FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1532
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)...(1294)
US-09-459-133-1

Query Match 4.4%; Score 44.8; DB 3; Length 1532;
Best Local Similarity 49.3%; Pred. No. 0.0022;
Matches 146; Conservative 0; Mismatches 147; Indels 3; Gaps 1;
Qy 247 CCCTTCCTCGTCTGCTGCTGACCTCATCCACAAACAGTTGGCTGAGCGCATGGCCATC 306
Db 542 CCTTACTGCTGTTGGCGTCAAGTCAGAACAGGGCGCTTTCGACAGACGACAGGCCGTG 601
Qy 307 CGGCAGACGTGGGGAAAGAGAGGACGGTGAAGGGAAGAGAGCTGAAGACATTTCTTC 366
Db 602 AGAGAGACGTGGGGCAGTCCAGCTCAGGGATCCGGCTCTTCTCTAGGGTCTCCG 661
Qy 367 CTGGGGACACACAGCAGTCCAGCGGAACAAAGAGTGGACAGGAGCAGCGACAC 426
Db 662 GTGGGTGAGCGGGCGCTGACCTAGACTCACTA---GTGGCATGGGAGAGCGCTCGCTAC 718
Qy 427 GGGGACATTTATCCAGAAGGATTTCTAGACGCTTATTACAATCTGACCCCTGAAGACCATG 486
Db 719 AGTGACCTGCTGCTCTGGGACTTCTCGACGTCCTCCATTCACACGACGCTCAAGACCTG 778
Qy 487 ATGGGCATAGAAATGGGTCCATCGCTTTTGTCTCAGCGGGCGTTTGTGATGAAC 542
Db 779 CTGCTGCTGGCTGGCTGGCGCGCACTGCCCCCGTGGTGTGCTTGTCTTGGAGC 834

RESULT 15
US-09-459-133-12
; Sequence 12, Application US/09459133
; Patent No. 6416988
; GENERAL INFORMATION:
; APPLICANT: Konklin, Darrell C.

; APPLICANT: Yamamoto, Gayle
; APPLICANT: Jaspers, Stephen R.
; APPLICANT: Gao, Zeren
; TITLE OF INVENTION: BETA-1,3-GALACTOSYLTRANSFERASE HOMOLOGS
; FILE REFERENCE: 98-77
; CURRENT APPLICATION NUMBER: US/09/459,133
; CURRENT FILING DATE: 1999-12-10
; PRIOR APPLICATION NUMBER: 60/111,697
; PRIOR FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 1446
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (107)...(1273)
US-09-459-133-12

Query Match 4.3%; Score 43.4; DB 3; Length 1446;
Best Local Similarity 49.1%; Pred. No. 0.006;
Matches 144; Conservative 0; Mismatches 146; Indels 3; Gaps 1;
Qy 247 CCCTTCCTCGTCTGCTGCTGACCTCATCCACAAACAGTTGGCTGAGCGCATGGCCATC 306
Db 524 CCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 583
Qy 307 CGGCAGACGTGGGGAAAGAGAGGACGGTGAAGGGAAGAGCTGAAGACATTTCTTC 366
Db 584 AGGAGACCTGGGGCAGCCAGTTGCTGGACCCGGTGTCTTCTCTGCTGGGTCCCCC 643
Qy 367 CTGGGGACACACAGCAGTCCAGCGGAACAAAGAGTGGACAGGAGAGAGCAGCGACAC 426
Db 644 CTAGG---AATGGGGGGCCTGACTTAAGATCACTGCTGACGTGGGAAAGCCGCGCTAT 700
Qy 427 GGGGACATTTATCCAGAAGGATTTCTTAGACGCTTATTACAATCTGACCCCTGAAGACCATG 486
Db 701 GGTGACCTTACTGCTCTGGGACTTCTCTGGATGTTCCCTTACAAACCGGACACTCAAGGACCTG 760
Qy 487 ATGGGCATAGAAATGGGTCCCATCGCTTTTGTCTCAGCGCGGCTTTGTGATGAA 539
Db 761 CTGCTGCTGACCTGGCTGAGCCACCACTGCCCGATGTCAATTTTGTCTTCTGCA 813

Search completed: September 22, 2005, 17:01:56
Job time : 213.477 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2005, 07:38:28 ; Search time 525.906 Seconds
(without alignments)
10595.706 Million cell updates/sec

Title: US-10-777-828-8_COPY_1_115

Perfect score: 115

Sequence: 1 ccactcagctcctcatgcat.....gagattgatgtatctgcc 115

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_on.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pt.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	115	100.0	1011	6	BD228354 UDP-galac
2	113.4	98.6	1360	9	AB041412 Gorilla g
3	113.4	98.6	1565	9	AB041413 Homo sapi
4	113.4	98.6	1576	9	AB041416 Homo sapi
5	113.4	98.6	10562	6	E38420 Novel polyp
6	113.4	98.6	170121	9	AF064860 Homo sapi
7	113.4	98.6	340000	9	HS21C080 Homo sapi
8	111.8	97.2	1566	9	AB041414 Pan trogl
9	111.8	97.2	1579	9	AB041417 Pongo pyg
10	111.8	97.2	192219	9	RP430021119 Pan trogl
11	110.2	95.8	1570	9	AB041415 Pan panis
12	58.2	50.6	170108	2	AC150794 Bos tauru
13	37.2	32.3	243047	2	AC126579 Rattus no
14	37	32.2	933	9	HGA6078 Homo sapi
15	36.8	32.0	131395	9	AC084740 Homo sapi
16	35.8	31.1	2494	6	CQ731786 Sequence
17	35.8	31.1	2762	9	AB020337 Homo sapi
18	35.8	31.1	2775	6	E38419 Novel polyp
19	35.8	31.1	2921	9	AY372061 Homo sapi

20	35.4	30.8	933	9	AF145784 Homo sapi
21	34.4	29.9	199196	2	AC130445 Rattus no
c 22	34.4	29.9	265998	2	AC127143 Rattus no
c 23	34.4	29.9	292664	2	AC103298 Rattus no
c 24	34.2	29.7	158959	2	AC122094 Rattus no
c 25	34.2	29.7	226295	2	AC119337 Rattus no
c 26	34	29.6	176814	2	AC121405 Rattus no
c 27	33.4	29.0	65608	9	AL450243 Human DNA
c 28	33.4	29.0	123229	2	AC102919 Mus muscu
c 29	33.4	29.0	181618	2	AL442129 Homo sapi
c 30	33.4	29.0	209860	2	AC151269 Mus muscu
c 31	33	28.7	78025	9	HS780M13 Human DNA
c 32	32.8	28.5	194018	2	AC138568 Gallus ga
c 33	32.2	28.0	1500	8	AB096162 Zostera c
c 34	32.2	28.0	161699	9	AC107300 Homo sapi
c 35	32.2	28.0	167357	2	AC026264 Homo sapi
c 36	32.2	28.0	168068	9	AL445668 Human DNA
c 37	32.2	28.0	191038	10	AC124719 Mus muscu
c 38	32	27.8	110001	9	AC112509 Homo sapi
c 39	32	27.8	167744	2	AC083941 Homo sapi
c 40	32	27.8	195269	6	AX417035 Sequence
c 41	32	27.8	195334	2	AC113657 Rattus no
c 42	32	27.8	200028	9	AC092680 Homo sapi
c 43	32	27.8	202354	9	AC006371 Homo sapi
c 44	32	27.8	333050	1	AL596168 Listeria
c 45	32	27.8	349980	6	AX417044 Sequence

ALIGNMENTS

RESULT 1	BD228354	1011 bp	DNA	linear	PAT 17-JUL-2003
LOCUS	UDP-galactose: beta-N-acetyl-glucosamine beta 1,3 galactosyltransferases, beta 3 Gal-T5.				
DEFINITION	BD228354				
ACCESSION	BD228354.1	GI:33038124			
VERSION	JP 2002530071-A/8.				
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 1011)				
AUTHORS	Clausen,H. and Amado,M.				
TITLE	UDP-galactose: beta-N-acetyl-glucosamine beta 1,3 galactosyltransferases, beta 3 Gal-T5				
JOURNAL	Patent: JP 2002530071-A 8 17-SEP-2002;				
COMMENT	HENRIK CLAUSEN				
	OS Homo sapiens (human)				
	PN JP 2002530071-A/8				
	PD 17-SEP-2002				
	PF 11-NOV-1999	JP 2000582542			
	PR 13-NOV-1998	DK PA 199801483			
	PI HENRIK CLAUSEN, MARGARIDA AMADO				
	PC C12N15/09, C12N1/15, C12N1/19, C12N5/10, C12N9/10, C12N15/00, C12N5/00				
	CC UDP-galactose: beta-N-acetyl-glucosamine beta 1,3 CC				
	galactosyltransferases,				
	CC beta 3 Gal-T5				
	Location/Qualifiers				
	FT Key				
	Location/Qualifiers				
	1..1011				
	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:9606"				

ORIGIN

Query Match 100.0%; Score 115; DB 6; Length 1011;
Best Local Similarity 100.0%; Pred. No. 5.6e-25;
Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

DEFINITION Homo sapiens betal,3-Galt 5 gene for UDP-Gal:GlcNAc
ACCESSION AB041416
VERSION AB041416.1 GI:7593026
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
REFERENCE 1 Homo sapiens
AUTHORS Narimatsu,H., Ieshiki,S., Togayauchi,A. and Sasaki,K.
TITLE Novel polypeptide
JOURNAL Patent: JP 2000245464-A 2 12-SEP-2000;
COMMENT KYOWA HAKKO KOGYO CO LTD
PN JP 2000245464-A/2
PD 12-SEP-2000
PF 25-FEB-1999 JP 1999047571
PR
PI HISASHI NARIMATSU,SOICHIRO ISSHIKI,AKIRA TOGAYAUCHI, PI
KATSUTOSHI SASAKI
PC C12N15/09,A01K67/027,C12N1/21,C12N5/10,C12N9/10,C12P19/00, PC
C12P21/02,
PC C12P21/08,C12Q1/68,G01N33/53//C12N1/21,C12R1:185) , (C12N5/10,
PC C12R1:91),
PC (C12P21/02,C12R1:185) , (C12P21/02,C12R1:91),C12N15/00,C12N5/00,
PC (C12N5/00,C12R1:91)
CC
FH Key Location/Qualifiers
FT promoter (1)..(5000)
FT exon (5001)..(5140)
FT exon (5001)..(5273)
FT exon (5459)..(5567)
FT exon (7427)..(7586)
FT exon (8234)..(10562).
FEATURES
source
1..10562
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 98.6%; Score 113.4; DB 6; Length 10562;
Best Local Similarity 99.1%; Pred. No. 1.9e-24;
Matches 114; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 CCACCTCAGCTCTCCTAGCATAAACTAGACATCCCTCATCTCTTTGAGGCTTAATCATT 60
Db 8156 CCACCTCAGCTCTCCTAGCATAAACTAGACATCCCTCATCTCTTTGAGGCTTAATCATT 8215
Qy 61 GGATTTTGTCTCTTCAGATGGCTTTCCGGAAGATGAGATTGATATATCTGCC 115
Db 8216 GGATTTTGTCTCTTCAGATGGCTTTCCGGAAGATGAGATTGATATATCTGCC 8270
RESULT 6
AF064860 170121 bp DNA linear PRI 05-MAR-2002
LOCUS Homo sapiens chromosome 21 clone PAC 70124 map 21q22.3, complete
DEFINITION
sequence.
ACCESSION AF064860
VERSION AF064860.2 GI:18958624
KEYWORDS HTG: HTGS DRAFT.
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 170121)
AUTHORS Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T.,
Park,H.-S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.-K., Soeda,E.,
Ohki,W., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K.,
Polley,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R.,
Zimmermann,W., Rosenthal,A., Kudoh,J., Kawasaki,K., Asakawa,S.,
Patterson,D., Reichwald,K., Rump,A., Schillhabel,M.B., Schudy,A.,
Shintani,A., Sasaki,T., Nagamine,K., Mitsuyma,S.,
Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordstiek,G.,
Hornischer,K., Brandt,P., Scharfe,M., Schoen,O., Desario,A.,
Heinzel,J., Kauer,G., Bloeker,H., Ramser,J., Beck,A., Klages,S.,
Hennig,S., RiesseImann,L., Dagand,E., Haaf,T., Wehrmeyer,S.,

```

```

DEFINITION Homo sapiens betal,3-Galt 5 gene for UDP-Gal:GlcNAc
ACCESSION AB041416
VERSION AB041416.1 GI:7593026
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
REFERENCE 1 Homo sapiens
AUTHORS Narimatsu,H., Ieshiki,S., Togayauchi,A. and Sasaki,K.
TITLE Novel polypeptide
JOURNAL Patent: JP 2000245464-A 2 12-SEP-2000;
COMMENT KYOWA HAKKO KOGYO CO LTD
PN JP 2000245464-A/2
PD 12-SEP-2000
PF 25-FEB-1999 JP 1999047571
PR
PI HISASHI NARIMATSU,SOICHIRO ISSHIKI,AKIRA TOGAYAUCHI, PI
KATSUTOSHI SASAKI
PC C12N15/09,A01K67/027,C12N1/21,C12N5/10,C12N9/10,C12P19/00, PC
C12P21/02,
PC C12P21/08,C12Q1/68,G01N33/53//C12N1/21,C12R1:185) , (C12N5/10,
PC C12R1:91),
PC (C12P21/02,C12R1:185) , (C12P21/02,C12R1:91),C12N15/00,C12N5/00,
PC (C12N5/00,C12R1:91)
CC
FH Key Location/Qualifiers
FT promoter (1)..(5000)
FT exon (5001)..(5140)
FT exon (5001)..(5273)
FT exon (5459)..(5567)
FT exon (7427)..(7586)
FT exon (8234)..(10562).
FEATURES
source
1..10562
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 98.6%; Score 113.4; DB 9; Length 1576;
Best Local Similarity 99.1%; Pred. No. 1.8e-24;
Matches 114; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 CCACCTCAGCTCTCCTAGCATAAACTAGACATCCCTCATCTCTTTGAGGCTTAATCATT 60
Db 597 CCACCTCAGCTCTCCTAGCATAAACTAGACATCCCTCATCTCTTTGAGGCTTAATCATT 656
Qy 61 GGATTTTGTCTCTTCAGATGGCTTTCCGGAAGATGAGATTGATATATCTGCC 115
Db 657 GGATTTTGTCTCTTCAGATGGCTTTCCGGAAGATGAGATTGATATATCTGCC 711
RESULT 5
E38420 10562 bp DNA linear PAT 31-JAN-2002
LOCUS Homo sapiens
DEFINITION Novel polypeptide.
ACCESSION E38420
VERSION E38420.1 GI:18626994
KEYWORDS JP 2000245464-A/2.
SOURCE Homo sapiens (human)

```


Borzym,K., Gardiner,K., Nizetic,D., Francis,F., Lehrach,H.,
Reinhardt,R. and Yaspo,M.Laure
The DNA sequence of human chromosome 21
Nature 405 (6784), 311-319 (2000)

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

2 (bases 1 to 170121)

Rump,A., Dagand,E., Hildmann,T., Nordsiek,G., Drescher,B.,
Weber,J., Schattevoy,R., Yaspo,M.-L. and Rosenthal,A.

TITLE
JOURNAL
REFERENCE
AUTHORS

Submitted (12-MAY-1998) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany

3 (bases 1 to 170121)

Rump,A., Dagand,E., Hildmann,T., Nordsiek,G., Drescher,B.,
Weber,J., Schattevoy,R., Yaspo,M.-L., Rosenthal,A., Yaspo,M.-L. and
Rosenthal,A.

TITLE
JOURNAL
REFERENCE
AUTHORS

Submitted (27-FEB-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany

4 (bases 1 to 170121)

Rump,A., Dagand,E., Hildmann,T., Nordsiek,G., Drescher,B.,
Weber,J., Schattevoy,R., Yaspo,M.-L. and Rosenthal,A.

TITLE
JOURNAL
COMMENT
FEATURES

Submitted (05-MAR-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
On Feb 27, 2002 this sequence version replaced gi:3171153.

Location/Qualifiers

1..170121
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
/clone="PAC 70124"

ORIGIN

Query Match 98.6%; Score 113.4; DB 9; Length 170121;
Best Local Similarity 99.1%; Pred.No.2.1e-24;
Matches 114; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CCACCTCAGCCTCCTAGCATAAACTAGACACATCCTCATGCTTTTGAGGTCCTAATCATTT 60

Db 89970 CCACCTCAGCCTCCTAGCATAAACTAGACACATCCTCATGCTTTTGAGGTCCTAATCATTT 90029

Qy 61 GGATTTTGTCTTTCAGATGGCTTTCCGAGATGAGATTCATGATATATCGCC 115

Db 90030 GGATTTTGTCTTTCAGATGGCTTTCCGAGATGAGATTCATGATATATTTGCC 90084

RESULT 7

HS21C080 340000 bp DNA linear PRI 24-MAY-2000
LOCUS Homo sapiens chromosome 21 segment HS21C080.
DEFINITION AL163280 AP001735 BA000005
ACCESSION AL163280.2 GI:7717369
VERSION

KEYWORDS
SOURCE
ORGANISM

Homo sapiens (human)

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 340000)

REFERENCE

AUTHORS
Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T.,
Park,H.S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.K., Soeda,E.,
Ohki,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K.,
Polley,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R.,
Patterson,D., Reichwald,K., Rump,A., Schillhabel,M., Schudy,A.,
Zimmermann,W., Rosenthal,A., Kudoh,J., Shibuya,K., Kawasaki,K.,
Asakawa,S., Shintani,A., Sasaki,T., Nagamine,K., Mitsuyama,S.,
Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordstieck,G.,
Hornischer,K., Brandt,P., Scharfe,M., Schoen,O., Desario,A.,
Reichelt,J., Kauer,G., Bloeker,H., Ramser,J., Beck,A., Klages,S.,
Hennig,S., Resselmann,L., Dagand,E., Wehrmeyer,S., Borzym,K.,
Gardiner,K., Nizetic,D., Francis,F., Lehrach,H., Reinhardt,R. and

Yaspo,M.L.

Direct Submission

Submitted (05-MAY-2000) The Chromosome 21 Mapping and Sequencing
Consortium: * RIKEN Genomic Sciences Center, Human Genome Research
Group * Institute of Molecular Biotechnology, Genome Analysis *
Keio University School of Medicine, Dept. of Molecular Biology *
GBF, Dept. of Genome Analysis * Max-Planck Institute for Molecular
Genetics (addresses see below)

COMMENT

The Chromosome 21 Mapping and Sequencing Consortium consists of
* RIKEN Genomic Sciences Center, Human Genome Research Group, *
Sagamihara 228-8555, Japan,
* e.mail: sakaki@gsc.riken.go.jp/
* URL: http://hgp.gsc.riken.go.jp/

and

* Institute of Molecular Biotechnology, Genome Analysis, *
Beutenbergstrasse 11, D-07745 Jena, Germany,
* e.mail: gscj-submit@genome.imb-jena.de
* URL: http://genome.imb-jena.de/

and

* Keio University School of Medicine, Dept. of Molecular Biology, *
Tokyo 160-8582, Japan,
* e.mail: shimizu@mb.med.keio.ac.jp
* URL: http://adenine.dmb.med.keio.ac.jp/

and

* GBF, Dept. of Genome Analysis,
* Mascheroder Weg 1, D-38124 Braunschweig, Germany, * e.mail:
info.genome@gbf.de
* URL: http://genome.gbf.de/

and

* Max-Planck Institute for Molecular Genetics,
* Ihnestrasse 73, D-14195 Berlin, Germany,
* e.mail: info-chr21@molgen.mpg.de
* URL: http://chr21.rz-berlin.mpg.de/.

FEATURES

source

Location/Qualifiers

1..340000
/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="21"

/map="21q22.3"

<1..125946

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="21"

/map="21q22.3"

/clone="P16019, 5' partial"

/clone_lib="RPC11,3-5 PAC library"

/note="Accession No. AF121897"

73410..243533

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="21"

/map="21q22.3"

/clone="P70124"

/clone_lib="RPC11,3-5 PAC library"

/note="Accession No. AF064860"

224137..2340000

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="21"

/map="21q22.3"

/clone="P206A10, 3' partial"

/clone_lib="RPC11,3-5 PAC library"

/note="Accession No. AF121782"

286628..2340000

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="21"

/map="21q22.3"


```

/mol_type="genomic DNA"
/isolate="chimp-202"
/db_xref="taxon:9598"
/ncfe="human sequence used for primer design based on Acc#
AB020337"
<1..28
/number=3
/number=3
/number=3
675..1566
/gene="betal,3-Galt 5"
675..1566
/gene="betal,3-Galt 5"
/codon_start=1
/product="UDP-Gal:GlcNAc betal,3-galactosyltransferase 5"
/db_xref="GI:7593023"
/translaton="MAPPKMLMTVCLLVLGALCYFMSYSLNLFKEQSFYKKDGNF
LKLPDTCRQTPPLVLLVTSRHLAERMAIRQTGWKRTVKGKQLTFLLGTTSS
AAETKEVDOSORHGDIIQKDLVDVYNYLTMTMGMEWVHRFCPOAAFYWKTDSDMF
INVDYLLKLNKRTTRFTGFLKLNELPIROPFSKFWKFSKSEYPWDRYPFPCSGTG
YVFGSVASQVYNVSESVYIKLEDVFGVGLCLERLNIRLEELHSQPTFFPGLRFSVC
RFRIRVACHFTKPTLLDYWQ"
675..1566
/gene="betal,3-Galt 5"
/number=4

ORIGIN
Query Match          97.2%; Score 111.8; DB 9; Length 1566;
Best Local Similarity 98.3%; Pred. No. 5.6e-24;
Matches 113; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCACCTCAGCCTCTAGCATAAACTAGACACATCCTCATGCTTTTGAGGCTCTAATCATT 60
    |||||
Db 597 CCACCTCAGCCTCTAGCATAAACTAGACACATCCTCATGCTTTTGAGGCTCTAATCATT 656

Qy 61 GGATTTTGTCTTTCAGATGGCTTCCGGAAGATGAGATTGATATATCTGCC 115
    |||||
Db 657 GGATTTTGTCTTTCAGATGGCTTCCGGAAGATGAGATTGATATATCTGCC 711

RESULT 9
AB041417
LOCUS          AB041417          1579 bp          DNA          linear          PRI 13-APR-2000
DEFINITION    Pongo pygmaeus betal,3-Galt 5 gene for UDP-Gal:GlcNAc
              betal,3-galactosyltransferase 5, partial cds.
ACCESSION     AB041417
VERSION       AB041417.1
KEYWORDS      'Pongo pygmaeus (orangutan)'
SOURCE        Pongo pygmaeus
ORGANISM      Pongo pygmaeus
REFERENCE     1 (bases 1 to 1579)
AUTHORS       Liu, Y. and Saitou, N.
TITLE         Silver Project
JOURNAL       Published Only in DataBase (2000)
REFERENCE     2 (bases 1 to 1579)
AUTHORS       Liu, Y. and Saitou, N.
TITLE         Direct Submission
JOURNAL       Submitted (11-APR-2000) Naruya Saitou, National Institute of
              Genetics, Laboratory of Evolutionary Genetics; 1111 Yata, Mishima,
              Shizuoka 411-8540, Japan (E-mail:nsaitou@genes.nig.ac.jp).
              URL: http://sayer.lab.nig.ac.jp/~silver/, Tel:81-559-81-6790,
              Fax:81-559-81-6789
              Location/Qualifiers
FEATURES             Location/Qualifiers
     source          1..1579
                     /organism="Pongo pygmaeus"
                     /mol_type="genomic DNA"
                     /isolate="oran-Po17"
                     /db_xref="taxon:9600"
                     /note="human sequence used for primer design based on Acc#
                     AB020337"

```

```

<1..26
/number=3
27..673
/number=3
674..1579
/gene="betal,3-Galt 5"
674..1579
/gene="betal,3-Galt 5"
/codon_start=1
/product="UDP-Gal:GlcNAc betal,3-galactosyltransferase 5"
/protein_id="BAA94502.1"
/db_xref="GI:7593023"
/translaton="MAPPKMLRIYICLLVLGALCLYFSMYNLPNPFKEQSFYKKEDGN
FLKLPDTCRQTPPLVLLVTSRHLAERMAIRQTGWKRTVKGKQLTFLLGTTSS
SAAEMKEVDOSORHGDIIQKDLVDVYNYLTMTMGMEWVHRFCPOAAFYWKTDSDM
FINVDYLLKLNKRTTRFTGFLKLNELPIROPFSKFWKFSKSEYPWDRYPFPCSGT
YVFGSVASQVYNVSESVYIKLEDVFGVGLCLERLNIRLEELHSQPTFFPGLRFSV
CRFRIVACHFTKPTLLDYWALEN"
674..1579
/gene="betal,3-Galt 5"
/number=4

ORIGIN
Query Match          97.2%; Score 111.8; DB 9; Length 1579;
Best Local Similarity 98.3%; Pred. No. 5.6e-24;
Matches 113; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCACCTCAGCCTCTAGCATAAACTAGACACATCCTCATGCTTTTGAGGCTCTAATCATT 60
    |||||
Db 596 CCACCTCAGCCTCTAGCATAAACTAGACACATCCTCATGCTTTTGAGGCTCTAATCATT 655

Qy 61 GGATTTTGTCTTTCAGATGGCTTCCGGAAGATGAGATTGATATATCTGCC 115
    |||||
Db 656 GGATTTTGTCTTTCAGATGGCTTCCGGAAGATGAGATTGATATATCTGCC 710

RESULT 10
RP43002119
LOCUS          RP43002119          192219 bp          DNA          linear          PRI 19-NOV-2003
DEFINITION    Pan troglodytes chromosome 22 BAC RP43-002119, complete sequence.
ACCESSION     AL954227
VERSION       AL954227.3
KEYWORDS      HTG.
SOURCE        Pan troglodytes (chimpanzee)
ORGANISM      Pan troglodytes
REFERENCE     1 (bases 1 to 192219)
AUTHORS       The Chimpanzee Chromosome 22 Sequencing Consortium.
TITLE         Chimpanzee chromosome 22 genomic sequence
JOURNAL       Unpublished
REFERENCE     2 (bases 1 to 192219)
AUTHORS       Scharte, M., Berg, C., Conrad, A., Hornischer, K., Loehnert, T.H.,
              Ludewig, M., Thies, S., Weber, K. and Bloecker, H.
TITLE         Direct Submission
JOURNAL       Submitted (25-JUN-2003) GBF, Dept. of Genome Analysis, Mascheroder
              Weg 1, D-38124 Braunschweig, Germany, E-mail: info.genome@gbf.de
              On Nov 19, 2003 this sequence version replaced gi:38228900.
              The Chimpanzee Chromosome 22 Sequencing Consortium consists of :
              *Chinese National Human Genome Center at Shanghai,
              Shanghai, China
              *GBF, Dept. of Genome Analysis, Braunschweig, Germany; *Institute
              of Molecular Biotechnology, Jena, Germany; *KRIBB Genome Research
              Center, Daejeon, Korea;
              *Max-Planck Institute for Molecular Genetics, Berlin, Germany;
              *National Institute of Genetics, Mishima, Japan;
              *National Yang Ming University Genome Research Center, Taipei,
              Taiwan;
              *RIKEN Genomic Sciences Center, Yokohama, Japan.
PROGRAMS AND PARAMETERS USED FOR ANNOTATION:
+++++
+ Analysis and annotation were performed with the automatic
+ 'first-pass' annotation and submission tool

```

```

+
+ 'AnnoMitter' (Hornischer & Bloecker).
+ Programs used by 'AnnoMitter':
+
+----- Genome Center
+ Center: GBF, Braunschweig
+ Center code: GBF
+ Web site: http://genome.gbf.de/
+ Contact: info.genome@gbf.de
+----- Project Information
+ Center project name:
+ Center clone name: RP43-002119
+----- Summary Statistics
+ Sequencing vector: ##;
+ Chemistry: Dye-terminator-amersham: ## of reads
+ Chemistry: Dye-primer-amersham: ## of reads
+ Assembly program: ##
+ Consensus quality: 0 bases at least Q40
+ Consensus quality: 0 bases at least Q30
+ Consensus quality: 0 bases at least Q20
+ Estimated insert size: ##; agarose-fp estimation
+ Estimated insert size: 192219; sum-of-contigs estimation
+-----
+
+ FEATURES
+ source
+ 1..192219
+ /organism="Pan troglodytes"
+ /mol_type="genomic DNA"
+ /db_xref="taxon:9598"
+ /chromosome="22"
+ /clone="RP43-002119"
+ 1..192219
+ /note="assembly fragment-clone end;T7-vector side:left
+ assembly_fragment-clone end;SP6-vector side:right"
+
+ misc_feature
+ 1..192219
+
+ ORIGIN
+ Query Match 97.2%; Score 111.8; DB 9; Length 192219;
+ Best Local Similarity 98.3%; Pred. No. 6.7e-24;
+ Matches 113; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
+
+ Qy 1 CCACCTGAGCTCTAGCATAAACTAGACATCCTCATGCTTTTGGTCTTAATCATT 60
+ Db 86593 CCACCTGAGCTCTAGCATAAACTAGACATCCTCATGCTTTTGGTCTTAATCATT 60
+
+ Qy 61 GGATTTTGTCTTTCAGATGCGCTTTCCGAGATGAGATGATGATATCTGCC 115
+ Db 86653 GGATTTTGTCTTTCAGATGCGCTTTCCGAGATGAGATGATGATGATGTTGCC 86707
+
+ RESULT 11
+ AB041415
+ LOCUS
+ DEFINITION
+ AB041415 Pan paniscus beta1,3-Galt 5 gene for UDP-Gal:GlcNAc
+ beta1,3-galactosyltransferase 5, partial cds.
+ ACCESSION
+ VERSION
+ AB041415.1 GI:7593024
+ SOURCE
+ Pan paniscus (pygmy chimpanzee)
+ ORGANISM
+ Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
+ Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
+ 1 (bases 1 to 1570)
+ TITLE
+ Silver Project
+ JOURNAL
+ Published Only in Database (2000)
+ 2 (bases 1 to 1570)
+ AUTHORS
+ Liu, Y. and Saitou, N.
+ DIRECT SUBMISSION
+ Submitted (11-APR-2000) Naruya Saitou, National Institute of
+ Genetics, Laboratory of Evolutionary Genetics; 1111 Yata, Mishima,
+ Shizuoka 411-8540, Japan (E-mail:nsaitou@genes.nig.ac.jp).
+ URL: http://sayer.lab.nig.ac.jp/~silver/, Tel: 81-559-81-6790,
+ Fax: 81-559-81-6789
+ FEATURES
+ Location/Qualifiers

```

```

1..1570
/organism="Pan paniscus"
/mol_type="genomic DNA"
/isolate="bonobo-05"
/db_xref="taxon:9597"
/note="human sequence used for primer design based on Acc#
AB020337"
exon <1..20
intron 21..667
gene 668..1570
CDS
668..>1570
/codon_start=1
/product="UDP-Gal:GlcNAc beta1,3-galactosyltransferase 5"
/protein_id="BAA94500.1"
/db_xref="GI:7593025"
/translation="MAFPKRLMYVCLLVGALCLYFSMYSNLNFKQSIFYKKDGNF
LKLPTDCROTTPPLVLLVTSKQLAERMAIROTGWKERTVQKQLKTFLLGTSS
AAETKVDOSORHGDIIQKDFLDGYNLTLKTMWGLEWHRFQPAAFVVKTDSDMF
INVVDYLLKKNRTTRFTGFLKNEPPIQPPSKFVSKVSPWDRYPDPCSTG
YVFSQVSVQVNVSVSPYIKLEDVFGVLCLERLNIRLELHSQPTFFPGGLRPSVC
RFRIVACHFIKPTLLDYQALEN"
exon 668..>1570
/gene="beta1,3-Galt 5"
/number=4
ORIGIN
Query Match 95.8%; Score 110.2; DB 9; Length 1570;
Best Local Similarity 97.4%; Pred. No. 1.7e-23;
Matches 112; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 CCACCTGAGCTCTAGCATAAACTAGACATCCTCATGCTTTTGGTCTTAATCATT 60
Db 590 CCACCTCAACCTCCTAGCATAAACTAGACATCCTCATGCTTTTGGTCTTAATCATT 649
Qy 61 GGATTTTGTCTTTCAGATGCGCTTTCCGAGATGAGATGATGATATCTGCC 115
Db 650 GGATTTTGTCTTTCAGATGCGCTTTCCGAGATGAGATGATGATGATGTTGCC 704
+
RESULT 12
AC150794 170108 bp DNA linear HTG 17-SEP-2004
LOCUS
DEFINITION
AC150794 Bos taurus clone CH240-397A7, WORKING DRAFT SEQUENCE, 6 ordered
pieces.
AC150794
VERSION
AC150794.3 GI:52219221
KEYWORDS
HTG; HTGS PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE
Bos taurus (cow)
ORGANISM
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovinae; Bos.
1 (bases 1 to 170108)
REFERENCE
Muzny, D., Adams, C., Agbai, I. O., Allen, C., Albrooks, S., Archer, P.,
Arronded, H., Bandaranaike, D., Bangura, L., Beltran, B., Beltran, R.,
Beraducci, A., Biswal, K., Blyth, P., Bonham, H., Buhay, C., Burch, P.,
Cadore, I., Canada, A., Cardenas, V., Carter, K., Cavazos, I.,
Chacko, J., Chahrour, M., Chavez, D., Chen, A., Chen, G., Chen, R.,
Cheng, M.-T., Chu, J., Clerc, K., Cockrell, R., Coyle, M., Cree, A.,
Curry, S., Dai, W., Davila, M. L., Davis, C., Davy-Carroll, L., De
Anda, C., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H.,
Donlin, J., McCauley, S., Dugan-Rocha, S., Dunn, A., Durbin, K.,
DiIuda, D., Egan, A., Escotto, M., Espinosa, V., Eugene, C., Fa, M.,
Fernandez, S., Fernando, P., Flagg, N., Forbes, L., Foster, P.,
Fowler, G., Fu, Q., Fuh, E., Garcia, A., Garcia, R., Garner, T.,
Gaskin, C., Gensch, S., Ghose, S., Gill, R., Gonzalez, D.,
Gonzalez-Garay, M., Guevara, W., Holder, M., Haaland, W., Haeblerlein, K.,
Hall, B., Hamid, H., Hamilton, K., Harbes, B., Harris, R., Haylak, P.,
Hawes, A., Hawkins, E., Hayes, S., Hemphill, L., Hernandez, J.,

```



```

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information -----
Center project name: GZJB
Center clone name: CH230-210M15
----- Summary Statistics -----
Assembly program: Phrap; version 0.990329
Consensus quality: 198187 bases at least Q40
Consensus quality: 201679 bases at least Q30
Consensus quality: 204254 bases at least Q20
Estimated insert size: 209786; sum-of-ctngs estimation
Quality coverage: 6x in Q20 bases; sum-of-ctngs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 243047: contig of 243047 bp in length.
* Location/Qualifiers
* 1..243047
* /organism="Rattus norvegicus"
* /mol_type="genomic DNA"
* /db_xref="taxon:10116"
* /clone="CH230-210M15"
* 1..1635
* /note="wgs_contig"

FEATURES
source
Query Match 32.3%; Score 37.2; DB 2; Length 243047;
Best Local Similarity 65.9%; Pred. No. 0.89;
Matches 54; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

ORIGIN
Qy 33 ATCTCATGCTTTTGAGGTCTAATCATTCGATTTTGTTCCTTTCAGATGCGCTTCCGAA 92
Db 107167 ATTTTCAGTATTAAAGAGATGATCATGTGATTTTTTTTCTTTTCTTTGTTGTGGCA 1071

```

Qy	93	GATGAGATTGATGTATATCTGC	114
Dd	107107	GATTATGTTGATGTATTTCTGC	107086

RESULT 14			
HSA6078			
LOCUS		933 bp	DNA linear PRI 11-MAY-2000
DEFINITION	Homo sapiens beta3gal-T6 gene.		
ACCESSION	AJ006078		
VERSION	AJ006078.1 GI:7799922		
KEYWORDS	beta-1,3-galactosyltransferase; beta3gal-T6 gene.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1	Anado,M., Carneiro,F. and Clausen,H. Cloning and expression of two beta-1,3-galactosyltransferases: beta3gal-T5 and beta3gal-T6 Unpublished	
AUTHORS		2 (bases 1 to 933) Anado,M.	
TITLE		Direct Submission	
JOURNAL		Submitted (11-MAY-1998) Amado M., Department of Oral Diagnostics, Royal dental School, Norre Alle 20, 2200 Copenhagen, DENMARK	
REFERENCE		Location/Qualifiers	
AUTHORS			
TITLE			
JOURNAL			
FEATURES			

```

source      1. 933
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
gene        1. 933
            /gene="beta3gal-T6"
CDS         1. 933
            /gene="beta3gal-T6"
            /codon_start=1
            /product="beta-1,3-galactosyltransferase"
            /protein_id="CAB91547.1"
            /db_xref="GI:7799923"
            /db_xref="GOA:Q9V2C3"
            /db_xref="UniProt/Swiss-Prot:Q9V2C3"
            /translation="MAFPMRLMYICLLVUGALCLYYSMSNLNPFKQSFVYKKGDNF
            LKLPDTCRQPPFLVLLVTSVSHKQLAERMAIQTWGKERTVVGKGLKTFLLGTTSS
            AAETKEVDQESRGDIIQKDFLDVYNNLTKTWNGIEWVHRFCFQAAFNKKTDSDMF
            INVLYELLLKKNVTRFFFLKLNFFIRQPFKWFVSKSPYMDRYPFPCSGTG
            YVSGDVASQVYVNSKSPYIKLEDVFVGLCLERLNIRLELHSHQPTFFPGGLAFSGC
            LFRIVACHFIKPRLLDYWQALENSRGEDCPV"

ORIGIN
Query Match      32.2%; Score 37; DB 9; Length 933;
Best Local Similarity 100.0%; Pred. No. 0.83;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 79 ATGGCTTCCGAGATGAGATTGATGATATCTGCCC 115
    |||||
Db 1 ATGGCTTCCGAGATGAGATTGATGATATCTGCCC 37

RESULT 15
AC084740
LOCUS      AC084740          131395 bp      DNA      linear      PRI 16-APR-2002
DEFINITION Homo sapiens BAC clone RP11-423E20 from 4, complete sequence.
ACCESSION AC084740
VERSION   AC084740.5 GI:19698727
KEYWORDS HTG.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 131395)
AUTHORS   Sulston, J.E. and Waterston, R.
TITLE     Toward a complete human genome sequence
JOURNAL   Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE   99063792
PUBMED    9847074
REFERENCE 2 (bases 1 to 131395)
AUTHORS   Levy, A. and Haakenson, W.
TITLE     The sequence of Homo sapiens BAC clone RP11-423E20
JOURNAL   Unpublished (2001)
REFERENCE 3 (bases 1 to 131395)
AUTHORS   Waterston, R.H.
TITLE     Direct Submission
JOURNAL   Submitted (10-NOV-2000) Genome Sequencing Center, Washington
           University School of Medicine, 4444 Forest Park Parkway, St. Louis,
           MO 63108, USA
REFERENCE 4 (bases 1 to 131395)
AUTHORS   Waterston, R.H.
TITLE     Direct Submission
JOURNAL   Submitted (14-MAR-2002) Genome Sequencing Center, Washington
           University School of Medicine, 4444 Forest Park Parkway, St. Louis,
           MO 63108, USA
REFERENCE 5 (bases 1 to 131395)
AUTHORS   Waterston, R.H.
TITLE     Direct Submission
JOURNAL   Submitted (24-MAR-2002) Genome Sequencing Center, Washington
           University School of Medicine, 4444 Forest Park Parkway, St. Louis,
           MO 63108, USA
REFERENCE 6 (bases 1 to 131395)
AUTHORS   Waterston, R.H.
TITLE     Direct Submission

```

```

JOURNAL     Submitted (29-MAR-2002) Genome Sequencing Center, Washington
           University School of Medicine, 4444 Forest Park Parkway, St. Louis,
           MO 63108, USA
REFERENCE   7 (bases 1 to 131395)
AUTHORS     Waterston, R.
TITLE       Direct Submission
JOURNAL     Submitted (16-APR-2002) Department of Genetics, Washington
           University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
           On Mar 24, 2002 this sequence version replaced gi:19424634.
COMMENT     ----- Genome Center
           Center: Washington University Genome Sequencing Center
           Center code: WUGSC
           Web site: http://genome.wustl.edu/gsc
           Contact: sapiens@watson.wustl.edu
           ----- Summary Statistics
           -----
           Center project name: H_NH0423E20
           -----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPC1-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Taten, M., Catanese, J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-312A15, 2000 bp overlap; the clone sequenced to the right is RP11-597D13, 2000 bp overlap. Actual start of this clone is at base position 114904 of RP11-312A15; actual end is at base position 6038 of RP11-597D13.

Polymorphisms exist between AC084740 and AC017037.

FEATURES

Location/Qualifiers

```

1. 131395
   /organism="Homo sapiens"
   /mol_type="genomic DNA"
   /db_xref="taxon:9606"
   /chromosome="4"
   /map="4"
   /clone="RP11-423E20"
   /clone_lib="RPC1-11"

```

repeat_region

```
1. .242
```

```
/rpt_family="Alu"
```

misc_feature

```
241. .819
   /note="similar to Homo sapiens EST AL534066"
```

```
(NID:gi2797559)"
```

misc_feature

```
241. .819
   /note="similar to Homo sapiens EST BE793889"
```

```
(NID:gi20215087)"
```

misc_feature

```
241. .819
   /note="similar to Homo sapiens EST BG675264"
```

```
misc_feature (NID:g13906660)"
241. .819 /note="similar to Homo sapiens EST BM006922
(NID:g16521276)"
misc_feature 241. .813 /note="similar to Homo sapiens EST BE619661
(NID:g9890599)"
misc_feature 241. .699 /note="similar to Homo sapiens EST BI856910
(NID:g15997657)"
misc_feature 241. .691 /note="similar to Homo sapiens EST BE173595
(NID:g8636321)"
misc_feature 241. .392 /note="similar to Homo sapiens EST BI868897
(NID:g16042570)"
misc_feature 241. .392 /note="similar to Homo sapiens EST BM019635
(NID:g16533989)"
misc_feature 241. .361 /note="similar to Homo sapiens EST AI042168 (NID:g3281362)
cy37b10.x1"
misc_feature 244. .361 /note="similar to Homo sapiens EST BG708944
(NID:g13986789)"
misc_feature 267. .369 /note="similar to Homo sapiens EST BE778176
(NID:g10199374)"
misc_feature 274. .819 /note="similar to Homo sapiens EST BE561211
(NID:g9804931)"
misc_feature 285. .819 /note="similar to Homo sapiens EST BG341756
(NID:g131148194)"
misc_feature 289. .819 /note="similar to Homo sapiens EST BM010255
(NID:g16524609)"
misc_feature 414. .923 /note="similar to EST BM171946 (NID:g17311509)"
misc_feature 446. .923 /note="similar to Bos taurus EST AV597915 (NID:g9715406)"
misc_feature 453. .819 /note="similar to Homo sapiens EST AW175759
(NID:g6441705)"
misc_feature 453. .819 /note="similar to Homo sapiens EST AW601923
(NID:g7306662)"
misc_feature 453. .819 /note="similar to Homo sapiens EST BG708944
(NID:g13986789)"
misc_feature 453. .814 /note="similar to Homo sapiens EST BE778176
(NID:g10199374)"
misc_feature 453. .814 /note="similar to Homo sapiens EST BM019635
(NID:g16533989)"
misc_feature 453. .797 /note="similar to Homo sapiens EST BE535420
(NID:g9764065)"
misc_feature 453. .485 /note="similar to Homo sapiens EST BE905008
(NID:g10397852)"
misc_feature 464. .819 /note="similar to Homo sapiens EST BI197251
(NID:g14652271)"
misc_feature 464. .804 /note="similar to Homo sapiens EST AI042168 (NID:g3281362)
cy37b10.x1"
misc_feature 464. .485 /note="similar to Homo sapiens EST BI868897
(NID:g16042570)"
misc_feature 541. .814 /note="similar to Homo sapiens EST BI868897
```

```
misc_feature (NID:g16042570)"
541. .756 /note="similar to Homo sapiens EST BE905008
(NID:g10397852)"
misc_feature 1179. 1202 /note="similar to Homo sapiens EST BG708944
(NID:g13986789)"
repeat_region 2333. .2585 /rpt_family="Alu"
repeat_region 3106. .3655 /rpt_family="MaLR"
repeat_region 3812. .3840 /rpt_family="AT-rich"
repeat_region 4393. .4702

Query Match 32.0%; Score 36.8; DB 9; Length 131395;
Best Local Similarity 61.5%; Pred. No. 1.2;
Matches 59; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

Qy 20 TAAAACTAGACACATCCTCATGCTTTTGAGGTCTAAATCATTGGATTTTGTCCCTTCAGA 79
Db 28052 TAAAAATTAGTCACATGGTCTTTTCATTGCAGGATGAATTTATTGAATGTGCTGCGCATGATCA 28111

Qy 80 TGGCTTTTCCCGAAGATGAGATTGATGTATATCTGCC 115
Db 28112 TTGCTAAACAGAAACTCAGAAATGGCCTACATCTGCC 28147

Search completed: September 22, 2005, 14:50:04
Job time : 529.906 secs
```

This Page Blank (uspto)

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	115	100.0	1011	3	AAA27959	Aaa27959 Human bet
2	113.4	98.6	596	12	ACH79926	Ach79926 Human gen
3	113.4	98.6	3409	10	ABZ56919	Abz56919 B3GALT nu
4	113.4	98.6	10562	3	AAa93876	Aaa93876 Human bet
5	37	32.2	933	10	ABZ56916	Abz56916 B3GALT nu
6	35.8	31.1	2629	10	ABZ56918	Abz56918 B3GALT nu
7	35.8	31.1	2762	10	ABZ56917	Abz56917 B3GALT nu
8	35.8	31.1	2775	3	AAa93875	Aaa93875 Human bet
9	35.4	30.8	933	12	ADK68456	Adk68456 Human bet
10	33	28.7	78025	8	ABQ77404	Abq77404 Human SEL
11	32	27.8	110000	6	ABQ69454_13	Continuation (14 o
12	32	27.8	110000	6	ABQ67195_3	Continuation (4 of
13	30.8	26.8	3796	6	ABQ67128	Abq67128 Human ang
14	30.6	26.6	2000	6	ABZ16396	Abz16396 Arabidops
15	30.4	26.4	2000	6	ABZ15338	Abz15338 Arabidops
16	30.4	26.4	335199	10	ADC24703	Adc24703 Human wil
17	30.2	26.3	260027	11	ACN44046	Acn44046 Human gen
18	30	26.1	99960	3	AAZ50905	Aaz50905 Human TBC
19	29.8	25.9	502	4	AAH11443	Aah11443 Human CDN
20	29.8	25.9	682	4	AAH33538	Aah33538 Human col

ALIGNMENTS

DR. P-PSDB; AAY9.

PT Novel nucleic acid sequence encoding human UDP-galactose:beta-N-
PT acetylglucosamine beta1,3-galactosyltransferase useful for obtaining beta
PT 1,3-galactosyl glycosylated saccharides and glycopeptides or
XX glycoproteins.

XX Claim 7; Fig 1; 74pp; English.

XX The present invention relates to a nucleic acid sequence encoding UDP-D-
CC galactose:beta-N-acetylglucosamine beta1,3-galactosyltransferase
CC (beta3Gal-T5). Beta3 transferases add galactose to the hydroxy group at
CC carbon 3 of 2-acetamido-2-deoxy-D-glucose (GlcNAc). The present sequence
CC represents the human beta3Gal-T5 gene sequence. The beta3Gal-T5 gene is
CC located on human chromosome 21q22.3. Beta3Gal-T5 is a type II
CC transmembrane glycoprotein. The invention also relates to the beta3Gal-T5
CC protein sequence, a nucleic acid vector comprising the beta3Gal-T5
CC nucleotide sequence, a host cell comprising the vector, and a method for
CC the production of the beta3Gal-T5 protein from the host cells. The
CC methods of the invention can be used for recombinant production of
CC beta3Gal-T5 for use as a catalyst and for recombinant production of
CC peptides or proteins with appropriate galactosylation. The beta3Gal-T5
CC protein can be used to obtain beta1,3-galactosyl glycosylated
CC saccharides, glycopeptides or glycoproteins

XX Sequence 1011 BP; 247 A; 256 C; 251 G; 257 T; 0 U; 0 Other;

Query Match 100.0%; Score 115; DB 3; Length 1011;
Best Local Similarity 100.0%; Pred. No. 6e-28; Length 1011;
Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCACCTCAGCCTCCTAGCATAAACTAGACACATCCTCATGCTTTGAGGCTTAATCATT 60
Db 1 CCACCTCAGCCTCCTAGCATAAACTAGACACATCCTCATGCTTTGAGGCTTAATCATT 60

Qy 61 GGATTTGTTCTTTCAGATGGCTTTCCGGAAGATGAGATTGATATATCTGCC 115
Db 61 GGATTTGTTCTTTCAGATGGCTTTCCGGAAGATGAGATTGATATATCTGCC 115

RESULT 2
ACH79926
ID ACH79926 standard; DNA; 596 BP.

XX ACH79926;

XX 29-JUL-2004 (first entry)

XX Human genome derived single exon probe #13121.

XX Human; probe; ss; gene expression; single exon probe; microarray;
XX alternative splicing event; genomic alteration.

XX Homo sapiens.

XX US2003194704-A1.

XX 16-OCT-2003.

XX 03-APR-2002; 2002US-00029386.

XX 03-APR-2002; 2002US-00029386.

XX (PENN/) PENN S G.
XX (RANK/) RANK D R.
XX (HANK/) HANKEL D K.

XX Penn SG, Rank DR, Hanzel DK;

XX WPI; 2004-119264/12.

XX New human genome-derived single exon nucleic acid probes useful for human
XX gene expression analysis, for identifying or characterizing alternative
XX splicing events, for assessing genomic alterations or as tools for
XX surveying tissues.

XX Claim 15; SEQ ID NO 13121; 80pp; English.

XX The invention relates to a nucleic acid probe for measuring human gene
CC expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridises under high stringency conditions to a nucleic acid molecule
CC expressed in human cells or tissues. Also included are a spatially-
CC addressable set of single exon nucleic acid probes for measuring human
CC gene expression (comprising a plurality of single exon nucleic acid
CC probes cited above, where each of the plurality of probes is separately
CC and addressably isolatable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of
CC measuring human gene expression, a vector comprising the single exon
CC probe cited above, an ORF-encoded peptide comprising at least 8
CC contiguous amino acids of any of the above-mentioned amino acid
CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above,
CC methods of selling and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, and a method of providing
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above. The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterising
CC alternative splicing events, in detecting and characterising gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030194704

XX Sequence 596 BP; 149 A; 150 C; 141 G; 156 T; 0 U; 0 Other;

Query Match 98.6%; Score 113.4; DB 12; Length 596;
Best Local Similarity 99.1%; Pred. No. 1.7e-27;
Matches 114; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CCACCTCAGCCTCCTAGCATAAACTAGACACATCCTCATGCTTTGAGGCTTAATCATT 60
Db 187 CCACCTCAGCCTCCTAGCATAAACTAGACACATCCTCATGCTTTGAGGCTTAATCATT 246

Qy 61 GGATTTGTTCTTTCAGATGGCTTTCCGGAAGATGAGATTGATATATCTGCC 115
Db 247 GGATTTGTTCTTTCAGATGGCTTTCCGGAAGATGAGATTGATATATCTGCC 301

RESULT 3

ABZ56919
ID ABZ56919 standard; DNA; 3409 BP.

XX AC ABZ56919;

XX 04-APR-2003 (first entry)

XX B3GALT nucleic acid sequence # SEQ ID 11.

XX B3GALT; beta-1,3 galactosyltransferase; p53; cancer; breast; colon;
XX kidney; lung; ovary; gene; ds.

XX Homo sapiens.

XX WO200299044-A2.

XX 12-DEC-2002.

PF 02-JUN-2002; 2002WO-US017356.
XX
PR 05-JUN-2001; 2001US-0296076P.
PR 10-OCT-2001; 2001US-0328605P.
PR 15-FEB-2002; 2002US-0357253P.
XX
PA (EXEL-) EXELIXIS INC.
XX
PI Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
XX WPI; 2003-156849/15.
XX
PT Identifying p53 pathway modulating agents with B3GALT genes, useful for
PT the diagnosis and treatment of disorders associated with defects in the
PT p53 pathway, such as cancer of the breast, colon, kidneys, lung and
XX ovary.
XX
PS Disclosure; Page 55-57; 82pp; English.
XX
XX The invention relates to identifying a candidate p53 pathway modulating
XX agent in humans that is referred to in the specification as B3GALT (beta-
XX 1,3 galactosyltransferase). Also disclosed is a method for diagnosing a
XX disease in a patient, by contacting a sample with a probe for B3GALT
XX expression, and comparing the results with a control, and determining
XX whether the results indicate a likelihood of disease. Methods and
XX compositions of the invention are useful for the diagnosis and treatment
XX of disorders associated with defects in the p53 pathway, such as cancer
XX of the breast, colon, kidneys, lung and ovary. The current sequence
XX represents a B3GALT nucleic acid sequence referred to in the disclosure
XX of the invention
XX
SQ Sequence 3409 BP; 833 A; 842 C; 833 G; 901 T; 0 U; 0 Other;
Query Match 98.6%; Score 113.4; DB 10; Length 3409;
Best Local Similarity 99.1%; Pred. No. 2.1e-27;
Matches 114; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CCACCTCAGCCTCCTAGCATAAACTAGACACATCCTCATGCTTTTGAGGTCCTAATCATT 60
DB 1003 CCACCTCAGCCTCCTAGCATAAACTAGACACATCCTCATGCTTTTGAGGTCCTAATCATT 1062
QY 61 GGATTTTGTCTTTTCAGATGCGTTTCCGAGATGAGATTGATATATCTGCC 115
DB 1063 GGATTTTGTCTTTTCAGATGCGTTTCCGAGATGAGATTGATATATCTGCC 1117
RESULT 4
AAA93876
ID AAA93876 standard; DNA; 10562 BP.
AC AAA93876;
XX
XX 15-JAN-2001 (first entry)
XX Human beta3Gal-T5 encoding DNA.
XX
XX Beta-1,3 galactose transferase; treatment; diagnosis; cancer; human;
XX digestive system; beta3gal-T5; ds.
XX
XX Homo sapiens.
XX
XX WO200050608-A1.
XX
XX 31-AUG-2000.
XX
XX 24-FEB-2000; 2000WO-JP001070.
XX
XX 25-FEB-1999; 99JP-00047571.
XX
XX (KYOWA) KYOWA HAKKO KOGYO KK.
XX
XX Narimatsu H, Isehiki S, Togayachi A, Sasaki K;
XX

DR WPI; 2000-549409/50.
XX
XX Beta-1,3 galactose transferase and DNA encoding it, useful for synthesis
PT of type 1 sialyl Lewis, a carbohydrate for treatment of digestive system
PT cancer.
XX
XX Claim 31; Page 103-111; 123pp; Japanese.
XX
XX This invention relates to a polypeptide (I) with beta-1,3 galactose
XX transferase activity, or variants of (I) comprising amino acid additions,
XX deletions and/or substitutions. Included in the invention is DNA encoding
XX all or part of (I); expression vectors containing the DNA, host cells
XX transformed by the vectors; a method for the preparation of the
XX polypeptide by culture of the transformants or by expression in the milk
XX of a transgenic mammal, and antibodies recognising (I). The Beta-1,3
XX galactose transferase protein transfers galactose by beta-1,3 bonding to
XX N-acetylglucosamine present in a non-cyclic carbohydrate chain (such as
XX GlcNAc-beta1-3gal-beta1-4Glc) to give gal-beta1-3GlcNAc. The protein and DNA
XX encoding it are useful for the treatment and diagnosis of cancer of the
XX digestive system. The present sequence represents a Beta3Gal-T5 encoding
XX DNA sequence
SQ Sequence 10562 BP; 2610 A; 2415 C; 2574 G; 2963 T; 0 U; 0 Other;
Query Match 98.6%; Score 113.4; DB 3; Length 10562;
Best Local Similarity 99.1%; Pred. No. 4.1e-27;
Matches 114; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CCACCTCAGCCTCCTAGCATAAACTAGACACATCCTCATGCTTTTGAGGTCCTAATCATT 60
DB 8156 CCACCTCAGCCTCCTAGCATAAACTAGACACATCCTCATGCTTTTGAGGTCCTAATCATT 8215
QY 61 GGATTTTGTCTTTTCAGATGCGTTTCCGAGATGAGATTGATATATCTGCC 115
DB 8216 GGATTTTGTCTTTTCAGATGCGTTTCCGAGATGAGATTGATATATCTGCC 8270
RESULT 5
ABZ56916
ID ABZ56916 standard; DNA; 933 BP.
XX
XX AC ABZ56916;
XX
XX 04-APR-2003 (first entry)
XX
XX B3GALT nucleic acid sequence # SEQ ID 8.
XX
XX B3GALT; beta-1,3 galactosyltransferase; p53; cancer; breast; colon;
XX kidney; lung; ovary; gene; ds.
XX
XX Homo sapiens.
XX
XX WO200299044-A2.
XX
XX 12-DEC-2002.
XX
XX 02-JUN-2002; 2002WO-US017356.
XX
XX 05-JUN-2001; 2001US-0296076P.
XX
XX 10-OCT-2001; 2001US-0328605P.
XX
XX 15-FEB-2002; 2002US-0357253P.
XX
XX (EXEL-) EXELIXIS INC.
XX
XX Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
XX WPI; 2003-156849/15.
XX
XX Identifying p53 pathway modulating agents with B3GALT genes, useful for
PT the diagnosis and treatment of disorders associated with defects in the
PT p53 pathway, such as cancer of the breast, colon, kidneys, lung and
XX ovary.
XX

PS Disclosure; Page 51-52; 82pp; English.

XX The invention relates to identifying a candidate p53 pathway modulating
CC agent in humans that is referred to in the specification as BGALT (beta-
CC 1,3 galactosyltransferase). Also disclosed is a method for diagnosing a
CC disease in a patient, by contacting a sample with a probe for B3GALT
CC expression, and comparing the results with a control, and determining
CC whether the results indicate a likelihood of disease. Methods and
CC compositions of the invention are useful for the diagnosis and treatment
CC of disorders associated with defects in the p53 pathway, such as cancer
CC of the breast, colon, kidneys, lung and ovary. The current sequence
CC represents a B3GALT nucleic acid sequence referred to in the disclosure
CC of the invention

XX Sequence 933 BP; 229 A; 234 C; 240 G; 229 T; 0 U; 1 Other;

Query Match 32.2%; Score 37; DB 10; Length 933;
Best Local Similarity 100.0%; Pred. No. 0.047;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 79 ATGGCTTTCCCGAGATGAGATTGATGTATATCTGCC 115
Db 1 ATGGCTTTCCCGAGATGAGATTGATGTATATCTGCC 37

RESULT 6

ABZ56918
ID ABZ56918 standard; DNA; 2629 BP.

AC ABZ56918;

DT 04-APR-2003 (first entry)

XX B3GALT nucleic acid sequence # SEQ ID 10.

XX B3GALT; beta-1,3 galactosyltransferase; p53; cancer; breast; colon;
KW kidney; lung; ovary; gene; ds.

XX Homo sapiens.

XX WO200299044-A2.

XX 12-DEC-2002.

XX 02-JUN-2002; 2002WO-US017356.

XX 05-JUN-2001; 2001US-0296076P.

PR 10-OCT-2001; 2001US-0328605P.

PR 15-FEB-2002; 2002US-0357253P.

XX (EXEL-) EXELIXIS INC.

XX Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;

XX WPI; 2003-156849/15.

XX Identifying p53 pathway modulating agents with B3GALT genes, useful for
PT the diagnosis and treatment of disorders associated with defects in the
PT p53 pathway, such as cancer of the breast, colon, kidneys, lung and
PT ovary.

PS Disclosure; Page 53-55; 82pp; English.

XX The invention relates to identifying a candidate p53 pathway modulating
CC agent in humans that is referred to in the specification as BGALT (beta-
CC 1,3 galactosyltransferase). Also disclosed is a method for diagnosing a
CC disease in a patient, by contacting a sample with a probe for B3GALT
CC expression, and comparing the results with a control, and determining
CC whether the results indicate a likelihood of disease. Methods and
CC compositions of the invention are useful for the diagnosis and treatment
CC of disorders associated with defects in the p53 pathway, such as cancer
CC of the breast, colon, kidneys, lung and ovary. The current sequence
CC represents a B3GALT nucleic acid sequence referred to in the disclosure

CC of the invention

XX Sequence 2629 BP; 632 A; 665 C; 644 G; 688 T; 0 U; 0 Other;

Query Match 31.1%; Score 35.8; DB 10; Length 2629;
Best Local Similarity 94.9%; Pred. No. 0.16;
Matches 37; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 77 AGATGGCTTTCCCGAGATGAGATTGATGTATATCTGCC 115
Db 299 AAATGGCTTTCCCGAGATGAGATTGATGTATATCTGCC 337

RESULT 7

ABZ56917
ID ABZ56917 standard; DNA; 2762 BP.

AC ABZ56917;

DT 04-APR-2003 (first entry)

XX B3GALT nucleic acid sequence # SEQ ID 9.

XX B3GALT; beta-1,3 galactosyltransferase; p53; cancer; breast; colon;
KW kidney; lung; ovary; gene; ds.

XX Homo sapiens.

XX WO200299044-A2.

XX 12-DEC-2002.

XX 02-JUN-2002; 2002WO-US017356.

XX 05-JUN-2001; 2001US-0296076P.

PR 10-OCT-2001; 2001US-0328605P.

PR 15-FEB-2002; 2002US-0357253P.

XX (EXEL-) EXELIXIS INC.

XX Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
XX WPI; 2003-156849/15.

XX Identifying p53 pathway modulating agents with B3GALT genes, useful for
PT the diagnosis and treatment of disorders associated with defects in the
PT p53 pathway, such as cancer of the breast, colon, kidneys, lung and
PT ovary.

PS Example 5; Page 52-53; 82pp; English.

XX The invention relates to identifying a candidate p53 pathway modulating
CC agent in humans that is referred to in the specification as BGALT (beta-
CC 1,3 galactosyltransferase). Also disclosed is a method for diagnosing a
CC disease in a patient, by contacting a sample with a probe for B3GALT
CC expression, and comparing the results with a control, and determining
CC whether the results indicate a likelihood of disease. Methods and
CC compositions of the invention are useful for the diagnosis and treatment
CC of disorders associated with defects in the p53 pathway, such as cancer
CC of the breast, colon, kidneys, lung and ovary. The current sequence
CC represents a B3GALT nucleic acid sequence referred to in an example from
CC the invention

XX Sequence 2762 BP; 651 A; 703 C; 679 G; 729 T; 0 U; 0 Other;

Query Match 31.1%; Score 35.8; DB 10; Length 2762;
Best Local Similarity 94.9%; Pred. No. 0.16;
Matches 37; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 77 AGATGGCTTTCCCGAGATGAGATTGATGTATATCTGCC 115
Db 432 AAATGGCTTTCCCGAGATGAGATTGATGTATATCTGCC 470

RESULT 8
 ID AAA93875 standard; DNA; 2775 BP.
 XX
 AC AAA93875;
 XX
 DT 15-JAN-2001 (first entry)
 XX
 DE Human beta-1,3 galactose transferase encoding DNA.
 XX
 KW Beta-1,3 galactose transferase; treatment; diagnosis; cancer; human;
 KW digestive system; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200050608-A1.
 XX
 PD 31-AUG-2000.
 XX
 PF 24-FEB-2000; 2000WO-JP001070.
 XX
 PR 25-FEB-1999; 99JP-00047571.
 XX
 PA (KYOM) KYOMA HAKKO KOGYO KK.
 XX
 PI Narimatsu H, Isshiki S, Togayachi A, Sasaki K;
 XX
 DR WPI; 2000-549409/50.
 DR P-PSDB; AAB93875.
 XX
 PT Beta-1,3 galactose transferase and DNA encoding it, useful for synthesis
 PT of type 1 sialyl Lewis, a carbohydrate for treatment of digestive system
 PT cancer.
 XX
 PS Claim 5; Page 99-102; 123pp; Japanese.
 XX
 CC This invention relates to a polypeptide (I) with beta-1,3 galactose
 CC transferase activity, or variants of (I) comprising amino acid additions,
 CC deletions and/or substitutions. Included in the invention is DNA encoding
 CC all or part of (I); expression vectors containing the DNA, host cells
 CC transformed by the vectors; a method for the preparation of the
 CC polypeptide by culture of the transformants or by expression in the milk
 CC of a transgenic mammal, and antibodies recognising (I). The Beta-1,3
 CC galactose transferase protein transfers galactose by beta-1,3 bonding to
 CC N-acetylglucosamine present in a non-cyclic carbohydrate chain (such as
 CC GlcNAc-beta1-3gal-beta1-4Glc) to give Gal-beta1-3GlcNAc. The protein and DNA
 CC encoding it are useful for the treatment and diagnosis of cancer of the
 CC digestive system. The present sequence represents Beta-1,3 galactose
 CC transferase encoding DNA
 XX
 SQ Sequence 2775 BP; 681 A; 698 C; 669 G; 727 T; 0 U; 0 Other;
 Query Match 31.1%; Score 35.8; DB 3; Length 2775;
 Best Local Similarity 94.9%; Pred. No. 0.16;
 Matches 37; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 77 AGATGGCTTCCCGAAGATGAGATTGATGATATCTGCC 115
 DB 400 AAATGGCTTCCCGAAGATGAGATTGATGATATCTGCC 438
 RESULT 9
 ADK68456
 ID ADK68456 standard; cDNA; 933 BP.
 XX
 AC ADK68456;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE Human beta-1,3-galactosyl transferase cDNA SeqID 1.
 XX
 KW human; gene; ss; saccharide binding protein; maltose binding protein;

KW beta-1,3-galactosyl transferase; inflammation; infectious disease;
 KW cancer metastasis suppression; dairy product; antiinflammatory;
 KW antimicrobial; cytostatic.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT CDS 1..933
 FT /*tag= a
 FT /product= "Beta-1,3-galactosyl transferase protein"
 XX
 PN JP2004016117-A.
 XX
 PD 22-JAN-2004.
 XX
 PF 17-JUN-2002; 2002JP-00176132.
 XX
 PR 17-JUN-2002; 2002JP-00176132.
 XX
 PA (TOYM) TOYOCO KK.
 XX
 DR WPI; 2004-113878/12.
 DR P-PSDB; ADK68457.
 XX
 PT New fusion protein useful in diagnosis of diseases and in manufacture of
 PT pharmaceutical products, comprises a saccharide binding protein and beta
 PT 1, 3-galactosyl transferase.
 XX
 PS Example 1; SEQ ID NO 1; 23pp; Japanese.
 XX
 CC This invention relates to a novel recombinant fusion protein that
 CC comprises a saccharide binding protein, in particular a maltose binding
 CC protein, and a beta-1,3-galactosyl transferase and an appropriate
 CC manufacturing method. Specifically, it refers to fusion protein that can
 CC transfer galactose to an N-acetyl glucosamine residue or N-acetyl
 CC glucosamine monosaccharide. The present invention describes a method to
 CC produce this enzyme, cheaply and efficiently, for the diagnosis and
 CC treatment of inflammation, infectious diseases or for cancer metastasis
 CC suppression. Furthermore, it can also be useful for improving dairy
 CC products. Accordingly, compositions exhibit various activities including
 CC antiinflammatory, antimicrobial and cytostatic. This polynucleotide
 CC sequence is the human beta-1,3-galactosyl transferase cDNA of the
 CC invention.
 XX
 SQ Sequence 933 BP; 227 A; 232 C; 241 G; 233 T; 0 U; 0 Other;
 Query Match 30.8%; Score 35.4; DB 12; Length 933;
 Best Local Similarity 97.3%; Pred. No. 0.16;
 Matches 36; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 79 ATGGCTTCCCGAAGATGAGATTGATGATATCTGCC 115
 DB 1 ATGGCTTCCCGAAGATGAGATTGATGATATCTGCC 37
 RESULT 10
 ABQ77404/c
 ID ABQ77404 standard; DNA; 78025 BP.
 XX
 AC ABQ77404;
 XX
 DT 10-MAY-2003 (first entry)
 XX
 DE Human SELP DNA.
 XX
 KW Human; SELP; vascular disease; cardiac; antiarteriosclerotic; stroke;
 KW cerebroprotective; gene therapy; coronary artery disease; ischaemia;
 KW myocardial infarction; peripheral vascular disease; pulmonary embolism;
 KW venous thromboembolism; forensic; paternity testing; GI3790157; gene;
 KW SNP; single nucleotide polymorphism; ds.
 XX
 OS Homo sapiens.
 XX

DT 28-AUG-2002 (first entry)
XX Human angiogenesis associated polynucleotide SEQ ID NO 158.
DE
XX
XX Human; angiogenesis; methylation; eye disease; glaucoma; tumour;
KW inflammation; rheumatoid arthritis; diabetic retinopathy; antiulcers;
KW macular degeneration; inflammatory bowel disease; Crohn's disease;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
XX antiarteriosclerotic; ds.
OS Homo sapiens.
XX
XX WO200246454-A2.
FN
XX
XX 13-JUN-2002.
PD
XX
XX 06-DEC-2001; 2001WO-EP014320.
PF
XX
XX 06-DEC-2000; 2000DE-01061338.
PR
XX
XX (EPIG-) EPIGENOMICS AG.
PA
XX
XX Schacht O;
PI
XX
XX WPI; 2002-500450/53.
DR
XX
XX New nucleic acid fragments from chemically treated angiogenesis-
PT associated genes, useful for determining methylation status, e.g. in
PT diagnosis or treatment of cancer.
PT
XX
XX Claim 1; SEQ ID NO 158; 41pp + Sequence Listing; German.
PS
XX
XX The invention relates to a nucleic acid (I) comprising a segment of 18
CC bases of chemically pretreated DNA of angiogenesis-associated genes (II)
CC having sequences (ABQ66971-ABQ67178) or their complements, (I)', also
CC related oligomers, are used to evaluate the methylation status and/or
CC single-nucleotide polymorphisms, in angiogenesis-related genes, for
CC diagnosis and treatment of eye diseases, proliferative retinopathy,
CC neovascular glaucoma, solid tumours, inflammation, rheumatoid arthritis,
CC diabetic retinopathy, macular degeneration caused by neovascularisation,
CC psoriasis, arteriosclerosis, inflammatory bowel diseases, ulcers and
CC Crohn's disease. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 3796 BP; 878 A; 151 C; 1102 G; 1665 T; 0 U; 0 Other;

Query Match 26.8%; Score 30.8; DB 6; Length 3796;
Best Local Similarity 61.0%; Pred. NO. 8.2;
Matches 50; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

Qy 31 ACATCCTCATGTTTGAGGTCATATCATCTTCTTCATGATTTCTTCATGATGCTTTCCG 90
Db 3311 AAATGTGAAGTAATGTAATTTATCTAGTAAATTTTATTTTAAAGATTTTAAATTG 3370

Qy 91 AAGATGAGATGATCATATATCT 112
Db 3371 GAGGTGGGATGATTTAAATAT 3392

RESULT 14
ABZ16396
ID ABZ16396 standard; DNA; 2000 BP.
XX
XX ABZ16396;
AC
XX
XX 21-JAN-2003 (first entry)
DT
XX
XX Arabidopsis thaliana stress regulated gene SEQ ID NO 4201.
DE
XX
KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX Arabidopsis thaliana.
OS

XX WO200216655-A2.
FN
XX
XX 28-FEB-2002.
PD
XX
XX 24-AUG-2001; 2001WO-US026685.
PF
XX
XX 24-AUG-2000; 2000US-0227866P.
PR
XX
XX 26-JAN-2001; 2001US-0264647P.
PR
XX
XX 22-JUN-2001; 2001US-0300111P.
PR
XX
XX (SCRI) SCRIPPS RES INST.
PA
XX
XX (SYGN) SYNGENTA PARTICIPATIONS AG.
FN
XX
XX Harper JF, Kreps J, Wang X, Zhu T;
PI
XX
XX WPI; 2002-304127/34.
DR
XX
XX Identifying a stress condition to which a plant cell has been exposed and
PT producing plants with increased tolerance to these abiotic stresses.
PT
XX
XX Claim 144; SEQ ID NO 4201; 577pp + Sequence Listing; English.
PS
XX
XX The invention relates to identifying a stress condition to which a plant
CC cell has been exposed, comprising: (a) contacting nucleic acid
CC representative of expressed polynucleotides in the plant cell with an
CC array or probes representative of the plant cell genome; and (b)
CC detecting a profile of expressed polynucleotides in the plant cell
CC characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants
CC with increased tolerance to abiotic stress. The present sequence is that
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used
CC in methods of the invention. Note: The sequence data for this patent is
CC not represented in the printed specification but is based on sequence
CC information supplied to Derwent by the European Patent Office
XX
SQ Sequence 2000 BP; 685 A; 313 C; 274 G; 728 T; 0 U; 0 Other;

Query Match 25.6%; Score 30.6; DB 6; Length 2000;
Best Local Similarity 55.0%; Pred. No. 7.9;
Matches 60; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

Qy 2 CACTCAGCTTCCTAGCATAAACTAGACATCTCTCATGCTTTTGGTCTTAATCATG 61
Db 1077 CACCAATAAATCAAGCACGGAACCTCGAAAGATCATGCTTTTCTGTCTAATTG 1136

Qy 62 GATTTGTTCTTTCAGATGCTTCCGAGATGAGATGATGATAT 110
Db 1137 TGGATTGAATTCATTTGGATTGGTTACCAAAAAAGAGAGATAATTT 1185

RESULT 15
ABZ15338/C
ID ABZ15338 standard; DNA; 2000 BP.
XX
XX ABZ15338;
AC
XX
XX 21-JAN-2003 (first entry)
DT
XX
XX Arabidopsis thaliana stress regulated gene SEQ ID NO 3143.
DE
XX
KW Arabidopsis thaliana; plant; gene; stress; transgenic; ds.
XX Arabidopsis thaliana.
OS
XX
XX WO200216655-A2.
FN
XX
XX 28-FEB-2002.
PD
XX
XX 24-AUG-2001; 2001WO-US026685.
PF
XX
XX 24-AUG-2000; 2000US-0227866P.
PR
XX
XX 26-JAN-2001; 2001US-0264647P.
PR


```

PR 22-JUN-2001; 2001US-0300111P.
XX
XX (SCEI ) SCRIPPS RES INST.
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
XX Harper JF, Krebs J, Wang X, Zhu T;
XX WPI; 2002-304127/34.
DR
XX
XX Identifying a stress condition to which a plant cell has been exposed and
PT producing plants with increased tolerance to these abiotic stresses.
XX
XX Claim 144; SEQ ID NO 3143; 577pp + Sequence Listing; English.
XX
XX The invention relates to identifying a stress condition to which a plant
CC cell has been exposed, comprising: (a) contacting nucleic acid
CC representative of expressed polynucleotides in the plant cell with an
CC array or probes representative of the plant cell genome; and (b)
CC detecting a profile of expressed polynucleotides in the plant cell
CC characteristic of a stress response. The method is useful in the
CC production of transgenic plants, cells and seeds and in producing plants
CC with increased tolerance to abiotic stress. The present sequence is that
CC of an Arabidopsis thaliana stress regulated gene (AB212196-AB217574) used
CC in methods of the invention. Note: The sequence data for this patent is
CC not represented in the printed specification but is based on sequence
CC information supplied to Derwent by the European Patent Office
XX
XX SQ Sequence 2000 BP; 563 A; 407 C; 385 G; 645 T; 0 U; 0 Other;

Query Match          26.4%; Score 30.4; DB 6; Length 2000;
Best Local Similarity 59.1%; Pred. No. 9.2;
Matches 52; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

Qy 19 ATAAACTAGACATCTCAGCTTTTGAGGCTCTAATCATTCATTGTTTCCTTCAG 78
Db 203 ATTTATCTGTCCTAATCTCATTTTACCTGGTTTGATCCTTGGGTTTGGCTCTCTAT 144

Qy 79 ATGGCTTTCCGAGATGAGATGATGT 106
Db 143 GTCGGCTTTTCGGAGATGGTTGATTT 116

```

Search completed: September 22, 2005, 12:38:51
Job time : 72.5965 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2005, 12:09:12 ; Search time 507.076 Seconds
(without alignments)
8632.613 Million cell updates/sec

Title: US-10-777-828-8_COPY_1_115

Perfect score: 115
Sequence: 1 ccacctcagctctcagcat.....gagattgatctatctgcc 115

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_ges1:*
9: gb_ges2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35.4	30.8	933	9	AY419460 Homo sapi
2	34	29.6	263	9	CB444598 tigr-ges-
3	33.8	29.4	780	9	AY419461 Pan trogl
4	33.4	29.0	848	8	AQ750178 HS_5575_A
5	32.8	28.5	436	6	CA834817 MCS031A04
6	32.6	28.3	879	5	B0395580 603804144
7	32.2	28.0	735	9	CR248164 Reverse s
8	31.8	27.7	701	8	BH932547 odh46f07.
9	31.8	27.7	839	9	CG457920 PUFYX23TB
10	31.8	27.7	889	9	CG169346 PUFY58TD
11	31.6	27.5	370	2	BE110064 UI-R-CA0-
12	31.4	27.3	685	2	BB275595 BB275595
13	31.2	27.1	424	7	CK515682 rswjbo_00
14	31.2	27.1	541	5	B0109768 imageq_7
15	31.2	27.1	584	4	BJ087028 BJ087028
16	31	27.0	492	7	CF272903 EST2465 Z
17	30.8	26.8	630	6	CB256589 70-E01079
18	30.8	26.8	760	2	BF794004 602254647
19	30.6	26.6	386	1	AU227311 AU227311
20	30.6	26.6	425	1	AU236398 AU236398
21	30.6	26.6	525	8	AQ370569 HS_5043_A
22	30.6	26.6	598	8	B2484243 BOOAB11TR
23	30.6	26.6	614	8	A2319319 1M0038C20
24	30.6	26.6	927	5	BX451934 BX451934

25	30.4	26.4	639	4	BJ488633
c 26	30.4	26.4	729	4	BJ748272
c 27	30.4	26.4	736	4	BJ509194
28	30.4	26.4	1174	6	CB561585 AGENCOURT
29	30.2	26.3	226	1	AV111150 AV111150
30	30.2	26.3	240	1	AJ484636 AJ484636
31	30.2	26.3	331	7	CO325225 EK195856.
32	30.2	26.3	360	1	AJ484637
c 33	30.2	26.3	779	8	BZ270094 CH230-275
34	30.2	26.3	1018	8	CC204655 CH261-152
35	30.2	26.3	1136	8	CC184602 CH261-820
36	30	26.1	541	9	CR076191 Forward s
c 37	30	26.1	633	9	CG107952 PUIJAS8TD
c 38	30	26.1	668	7	CV472267
c 39	30	26.1	742	9	AG414481 Mus muscu
c 40	30	26.1	792	8	BH951588 odh16b07.
41	30	26.1	820	9	CR024687 Forward s
c 42	29.8	25.9	224	1	AA810738
c 43	29.8	25.9	235	1	A1004187
c 44	29.8	25.9	266	1	A1524537
c 45	29.8	25.9	347	1	AA824559

ALIGNMENTS

RESULT 1
LOCUS AY419460 933 bp DNA linear GSS 17-DEC-2003
DEFINITION Homo sapiens B3GALT5 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION AY419460
VERSION AY419460.1 GI:39775417
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 933)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 933)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.
TITLE Direct Submision
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT These sequences were made by sequencing genomic exons and ordering them based on alignment.
FEATURES
Location/Qualifiers
1..933
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>933
/gene="B3GALT5"
/locus_tag="HCM6898"

Query Match 30.8%; Score 35.4; DB 9; Length 933;
Best Local Similarity 97.3%; Pred. No. 3.5;
Matches 36; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 79 ATGCGTTTCCGAGATGAGATTGATGATATCTGCC 115
|||||

1 ATGCGTTTCCGGAAGATGAGATTGATGATATTGTC 37

CE444598 263 bp DNA linear GSS 27-SEP-2003
tigr-gss-dog-17000336367589 Dog Library Canis familiaris genomic,
genomic survey sequence.
CE444598
CE444598.1 GI:36729674
GSS.
Canis familiaris (dog)
Canis familiaris
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
1 (bases 1 to 263)
Kirkness,E.F., Bafna,V., Halpern,A.L., Levy,S., Remington,K.,
Rusch,D.B., Deicher,A.L., Pop,M., Wang,W., Fraser,C.M. and
Venter,J.C.
The dog genome: survey sequencing and comparative analysis
Science 301 (5641), 1898-1903 (2003)
22875432
14512627
Contact: Kirkness EF
The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirknes@tigr.org
Class: shotgun.
Location/Qualifiers
1..263
/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/notes="Site 1: BatXI; Libraries were prepared from
peripheral blood"

ORIGIN
Query Match 29.6%; Score 34; DB 9; Length 263;
Best Local Similarity 63.4%; Pred. No. 7.8;
Matches 52; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
Qy 27 AGACACATCCTCAGCTTTTGAGTCTAATCATGGAATTTTGTCCTTCAGATGGCTTT 86
Db 1 AAACCGGTCAATCATGTTCCTTCGCTCTTTTCGTTCCATGTTGTTCCCTTAGTTCCTGTT 60
Qy 87 CCCGAGATGAGATTGATGTAT 108
Db 61 GTTTAAGATTTTATTTAIGTAT 82

RESULT 3
AY419461
LOCUS
DEFINITION
Pan troglodytes B3GAL75 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
AY419461
AY419461.1 GI:39775418
GSS.
Pan troglodytes (chimpanzee)
Pan troglodytes
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Pan.
1 (bases 1 to 780)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Forriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous

ORIGIN

Query Match	28.3%;	Score 32.6;	DB 5;	Length 879;
Best Local Similarity	58.9%;	Pred. No. 26;		
Matches 56;	Conservative 0;	Mismatches 39;	Indels 0;	Gaps 0;

```

Qy 11 CTCCTGATCAAAACTAGACACATCTCATGCTTTTGAGGTCTAATCATTTGGATTGTT 70
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 292 CTCGTATAAGACAACCTGCACCTGAGCATTAATGGTCTATAGATCATGCTCTTAGCTTAAGTA 351
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 71 CTTTTCAGATGGCTTTCCCGAAGATGAGATTGATG 105
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 352 CCTAATACTTAGCAGTGCATCAGGTGAGATTGATG 386
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 7
CR248164 735 bp DNA linear GSS 06-JUL-2004
LOCUS Reverse strand read from insert in 5'HPRT insertion targeting and
DEFINITION chromosome engineering clone MHPN85b22, genomic survey sequence.
ACCESSION CR248164
VERSION CR248164.1 GI:50027018
KEYWORDS GSS; genome survey sequence; MICR.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 735)
AUTHORS Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
Jongers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
Rogers,J. and Bradley,A.
Direct Submission
Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. http://www.sanger.ac.uk/MICR
JOURNAL Location/Qualifiers
FEATURES
    source
        1..735
            /organism="Mus musculus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10090"
            /clone="MHPN85b22"
            /clone_lib="MHPN"

ORIGIN
Query Match 28.0%; Score 32.2; DB 9; Length 735;
Best Local Similarity 57.4%; Pred. NO. 34;
Matches 58; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

Qy 15 TAGCATAAACTAGACACATCTCATGCTTTTGAGGTCTAATCATTTGGATTGTTGCTCT 74
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 443 TAAGAAATGCTCTACTCTATATATTAATGAGTCTTTCTCAATGGGTTTCTCTCTT 502
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 75 TCAGATGGCTTTCCCGAAGATGAGATTGATGATATCTGCC 115
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 503 TCAGATGACTTAGCTAATATATCATGTTGCACATGAACTGTC 543
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 8
BH932547 701 bp DNA linear GSS 01-OCT-2002
LOCUS osh46f07.g1 B.oleracea002 Brassica oleracea genomic, genomic survey
DEFINITION sequence.
ACCESSION BH932547
VERSION BH932547.1 GI:23412613
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 701)
AUTHORS Delehaunty,K., Fewell,G., Fulton,L., McCombie,W.R., Miner,T.,
Nash,W., Rabinowicz,P.D. and Wilson,R.K.
Whole genome shotgun reads from Brassica oleracea
Contact: Richard K. Wilson
Unpublished (2002)
JOURNAL Genome Sequencing Center
COMMENT Washington University School of Medicine
Email: submissions@watson.wustl.edu
Plate: osh46 row: f column: 07

```

```

Seq primer: -28RPpOT reverse
Class: shotgun
High quality sequence start: 16
High quality sequence stop: 217.
FEATURES
    Location/Qualifiers
        1..701
            /organism="Brassica oleracea"
            /mol_type="genomic DNA"
            /db_xref="taxon:3712"
            /clone_lib="B.oleracea002"
            /note="vector: pOTw13; Whole genome shotgun library from
            flowering buds. DNA was purified from a crude nuclear
            prep using Brassica oleracea T01000DH3 buds provided by
            Thomas Osborn at the University of Wisconsin. Genomic
            DNA was provided by Pablo Rabinowicz (CSHL) and the
            shotgun library prepared at Washington University Genome
            Sequencing Center."

ORIGIN
Query Match 27.7%; Score 31.8; DB 8; Length 701;
Best Local Similarity 64.0%; Pred. No. 44;
Matches 48; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

Qy 14 CTAGCATAAACTAGACACATCTCATGCTTTTGAGGTCTAATCATTTGGATTGTTGCT 73
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 315 CGATTATGAACTTTTAACATCTTCATCTGTTGGGTATGTTATTGTCCTTGGTGCT 374
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 74 TTCAGATGGCTTTCC 88
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 375 TGATGAAGTGTTC 389
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 9
CG457920 839 bp DNA linear GSS 17-SEP-2003
LOCUS PUFXM23TBC ZM 0.6 1.0 KB Zea mays genomic clone ZMMB7a0755D21,
DEFINITION genomic survey sequence.
ACCESSION CG457920
VERSION CG457920.1 GI:34842920
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 839)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
Maize Genomics Consortium
Unpublished (2003)
JOURNAL Other GSSs: PUFXM23TDC
COMMENT Contact: Cathy Whitelaw
TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
FEATURES
    Location/Qualifiers
        1..839
            /organism="Zea mays"
            /mol_type="genomic DNA"
            /strain="B73"
            /db_xref="taxon:4577"
            /clone="ZMMB7a0755D21"
            /clone_lib="ZM 0.6 1.0 KB"
            /note="vector: PCR4-TOPO; site_1: EcoRI; 0.6-1.0 kb high
            Cor selected genomic DNA library"

ORIGIN
Query Match 27.7%; Score 31.8; DB 9; Length 839;

```

```

Best Local Similarity 56.1%; Pred. No. 46;
Matches 60; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

Qy 6 TCAGCCTCCTAGCATAAACATAGACACATCCTCATGCTTTTGAGGTCTAATCATTTGGATT 65
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 31 TGAAGTTATATGTAATCCATTCATTAACCTTAGGTTAGAAACACATATCATTTCAATT 90
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 66 TTGTTCCCTTCAGATGGCTTTCCCGAAGATGAGATTGATGATATCT 112
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 91 TTATTCCCTTTAGGATAAACTAAGTCATATATATATATATATATATATATATATATATAT 137
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10
CG169346 889 bp DNA linear GSS 21-AUG-2003
PUJPE56TD ZM 0.6-1.0 KB Zea mays genomic clone ZMWBTA0654J15,
genomic survey sequence.
ACCESSION CG169346
VERSION CG169346.1 GI:34060147
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 889)
White, A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and
Bennetzen, J.
Maize Genomics Consortium
Unpublished (2003)
Other GSSs: PUJPE56TB
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TP
Class: sheared ends.

FEATURES
    source
        1..889
            /organism="Zea mays"
            /mol_type="genomic DNA"
            /strain="B73"
            /db_xref="taxon:4577"
            /clone="ZMWBTA0654J15"
            /note="Vector: PCR4-TOPO; site_1: EcoRI; 0.6-1.0 kb high
            Cot selected genomic DNA library"

ORIGIN
Query Match 27.7%; Score 31.8; DB 9; Length 889;
Best Local Similarity 56.1%; Pred. No. 46;
Matches 60; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

Qy 6 TCAGCCTCCTAGCATAAACATAGACACATCCTCATGCTTTTGAGGTCTAATCATTTGGATT 65
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 320 TGAAGTTATATGTAATCCATTCATTAACCTTAGGTTAGAAACACATATCATTTCAATT 379
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 66 TTGTTCCCTTCAGATGGCTTTCCCGAAGATGAGATTGATGATATCT 112
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 380 TTATTCCCTTTAGGATAAACTAAGTCATATATATATATATATATATATATATATATATAT 426
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 11
BE110064/c 370 bp mRNA linear EST 13-JUN-2000
LOCUS BE110064
DEFINITION UI-R-CA0-axj-f-03-0-UI-s1 UI-R-CA0 Rattus norvegicus cDNA clone
ACCESSION BE110064
VERSION BE110064.1 GI:8502169
KEYWORDS EST.

```

```

SOURCE
ORGANISM Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 370)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
889548
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to identify it as a clone from the
normalized hippocampus library cDNA Library Preparation: M.B.
Soares Lab Clone distribution: clones will be available through
Research Genetics (www.resgen.com)
Seq primer: M13 Forward
POLYA-Yes.
    Location/Qualifiers
        1..370
            /organism="Rattus norvegicus"
            /mol_type="mRNA"
            /strain="Sprague-Dawley"
            /db_xref="taxon:10116"
            /clone="UI-R-CA0-axj-f-03-0-UI"
            /lab_host="DH10B (Life Technologies)"
            /clone_lib="UI-R-CA0"
            /note="Vector: p7T3D-Pac (Pharmacia) with a modified
            polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-CA0
            library is a subtracted library derived from the following
            tissues: thalamus, cerebellum, hypothalamus, medulla,
            pons, midbrain, cerebral cortex, corpus striatum, testis,
            and hippocampus. For a detailed description of the
            library from which this clone was derived, please visit
            our web site at ratest.eng.uiowa.edu. The subtraction
            has been previously described in (Bonaldo, Lennon and
            Soares, Genome Research 6:791-806, 1996)
            TAG LIB=UI-R-CA0
            TAG_SEQ=GATTG"

ORIGIN
Query Match 27.5%; Score 31.6; DB 2; Length 370;
Best Local Similarity 60.5%; Pred. No. 46;
Matches 52; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 7 CAGCCTCCTAGCATAAACATAGACACATCCTCATGCTTTTGAGGTCTAATCATTTGGATT 66
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 128 CTGCCATCTAGCACAAGACTGGGTACATCTCTCATGTGTGGATTTCTCTTGTGGCTTC 69
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 67 TGTTCCCTTCAGATGGCTTTCCCGAA 92
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 68 ATTTCTTACTCATGGACTCCCGAA 43
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 12
BB275595 685 bp mRNA linear EST 24-OCT-2001
LOCUS BB275595
DEFINITION BB275595 RIKEN full-length enriched, 10 days neonate cortex Mus
musculus cDNA clone AB3009JM07 3', mRNA sequence.
ACCESSION BB275595
VERSION BB275595.2 GI:16400850

```


QY 15 TAGCATAAACTAGACACATCTCATGCTTTTGAGGTCTAATCATGATTTGTTCCCTT 74
|||||
Db 375 TAGCGTACATCTCGAATCTTTTACAGGTTCTGGAATCCAAACATTTGTTTGAATCGTT 316
|||||
QY 75 TCAGATGCGCTTTCCCGAAGATGAGATTGATG 106
|||||
Db 315 TCGCCCACTGACCAAGATTGAGGTGATTT 284
|||||

RESULT 14
LOCUS BQ109768/c 541 bp mRNA linear EST 16-APR-2002
DEFINITION imageqc_7_2001/snn49bdf41.xl NICHG_XGC_Embi Xenopus laevis cDNA
clone IMAGE:5161929 3', mRNA sequence.
ACCESSION BQ109768
VERSION BQ109768.1 GI:20159422
KEYWORDS EST.
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Xenopus laevis
REFERENCE 1 (bases 1 to 541)
AUTHORS Kale, P.I., Harsch, T.J., Folta, P.A., Nelson, D.O., Sanders, C.G. and Prange, C.K.
TITLE The I.M.A.G.E. Consortium quality control effort: clone
resequencing for verification
JOURNAL Unpublished (2001)
COMMENT Contact: Prange CK
The I.M.A.G.E. Consortium
Lawrence Livermore National Laboratory
Livermore, CA, USA
Email: help@image.llnl.gov
This read has been produced as part of the I.M.A.G.E. Consortium
quality control effort. High quality sequence is defined as having
100 or more base pairs with a phred quality value of 20 or greater,
where a sliding window of 4 base pairs with a phred quality value
of 15 or greater marks the beginning and end of the sequence. For
information on obtaining this clone, please contact
info@image.llnl.gov.
Plate: LLAM11401 row: n column: 10
Seq primer: -21ml3
High quality sequence stop: 541.
Location/Qualifiers
1..541
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE:5161929"
/tissue type="embryo (stage 10)"
/lab host="DH10B (phage-resistant)"
/clone lib="NICHG_XGC_Emb1"
/note="Vector: PCMV-SF0RT6; Site 1: NotI; Site 2: SalI;
Cloned unidirectionally. Primer: Oligo dt. Average insert
size 1.55 kb. Constructed by Life Technologies. Note: This
is a Xenopus Gene Collection (XGC) library."

ORIGIN

Query Match 27.1%; Score 31.2; DB 5; Length 541;
Best Local Similarity 58.7%; Pred. No. 65;
Matches 54; Conservative 0; Mismatches 38; Indels 0; Gaps 0;
QY 18 CATAAACTAGACACATCTCATGCTTTTGAGGTCTAATCATGATTTGTTCCCTTCA 77
|||||
Db 442 CATAGATTATTACAATACTAGTGCTTTCAGGATTGAATTAGATTTTTTTTGCTT 383
|||||
QY 78 GATGCTTTCCCGAAGATGAGATTGATGTATA 109
|||||
Db 382 AATAGCATTTACGGAAATAGATTATGTATA 351
|||||

RESULT 15
BQ087028/c

LOCUS BQ087028 584 bp mRNA linear EST 29-SEP-2003
DEFINITION BQ087028 NIBB Mochii normalized Xenopus tailbud library Xenopus
laevis cDNA clone XL055d11 3', mRNA sequence.
ACCESSION BQ087028
VERSION BQ087028.1 GI:17583656
KEYWORDS EST.
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Xenopus laevis
REFERENCE 1 (bases 1 to 584)
AUTHORS Kitayama, A., Terasaka, C., Mochii, M., Ueno, N., Shin-i, T. and Kohara, Y.
TITLE Expressed genes in X. laevis embryo
JOURNAL Unpublished (2001)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp
The information of this clone is available through the following
URL.
http://xenopus.nibb.ac.jp.
Location/Qualifiers
1..584
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="XL055d11"
/tissue type="whole embryo"
/dev stage="stage 25"
/clone lib="NIBB Mochii normalized Xenopus tailbud
library"

FEATURES
source

ORIGIN

Query Match 27.1%; Score 31.2; DB 4; Length 584;
Best Local Similarity 58.7%; Pred. No. 66;
Matches 54; Conservative 0; Mismatches 38; Indels 0; Gaps 0;
QY 18 CATAAACTAGACACATCTCATGCTTTTGAGGTCTAATCATGATTTGTTCCCTTCA 77
|||||
Db 430 CATAGATTATTACAATACTAGTGCTTTCAGGATTGAATTAGATTTTTTTTGCTT 371
|||||
QY 78 GATGCTTTCCCGAAGATGAGATTGATGTATA 109
|||||
Db 370 AATAGCATTTACGGAAATAGATTATGTATA 339
|||||

Search completed: September 22, 2005, 16:56:02
Job time : 516.076 secs

his Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2005, 12:15:17 ; Search time 23.9415 Seconds
(without alignments)
7859.644 Million cell updates/sec

Title: US-10-777-828-8_COPY_1_115

Perfect score: 115

Sequence: 1 ccacctcagctcctcctagcat.....gagattgatctatctgcc 115

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	115	100.0	1011	4	US-09-831-630-8
2	30.2	26.3	327	4	US-09-270-767-2980
3	30.2	26.3	327	4	US-09-270-767-18262
4	30	26.1	99960	4	US-09-762-311-2
5	29	25.2	359	4	US-08-958-171E-2970
6	29	25.2	359	4	US-08-781-986A-2970
7	28.8	25.0	2330	3	US-09-120-653D-4
8	28.8	25.0	3910	3	US-09-120-653D-1
9	28.4	24.7	524032	4	US-09-949-016-16928
10	28.4	24.7	524032	4	US-09-949-016-16929
11	28.4	24.7	524032	4	US-09-949-016-16930
12	28.4	24.7	524032	4	US-09-949-016-16931
13	28.4	24.7	529885	4	US-09-949-016-14340
14	28.4	24.7	529885	4	US-09-949-016-14341
15	28.4	24.7	529885	4	US-09-949-016-14342
16	28.4	24.7	529885	4	US-09-949-016-14343
17	28.4	24.7	529885	4	US-09-949-016-14344
18	28.4	24.7	529885	4	US-09-949-016-14345
19	28.4	24.7	529885	4	US-09-949-016-14346
20	28.4	24.7	529885	4	US-09-949-016-14347
21	28.2	24.5	68490	4	US-09-949-016-15849
22	28.2	24.2	10321	4	US-09-949-016-13587
23	27.6	24.0	601	4	US-09-949-016-149681
24	27.6	24.0	640681	1	Sequence 1, Appli
25	27.4	23.8	513	3	US-09-188-930-237
26	27.4	23.8	513	4	US-09-312-283C-237
27	27.4	23.8	2124	4	US-09-328-352-708

28	27.4	23.8	40936	4	US-09-949-016-16607	Sequence 16607, A
29	27.4	23.8	40936	4	US-09-949-016-16608	Sequence 16608, A
c 30	27.4	23.8	205163	4	US-09-949-016-17009	Sequence 17009, A
c 31	27.2	23.7	601	4	US-09-949-016-51191	Sequence 51191, A
32	27.2	23.7	17226	4	US-09-949-016-13806	Sequence 13806, A
c 33	27.2	23.7	118923	4	US-09-949-016-13227	Sequence 13227, A
c 34	27.2	23.7	153866	4	US-09-949-016-16919	Sequence 16919, A
c 35	27.2	23.7	235064	4	US-09-949-016-15390	Sequence 15390, A
36	27.2	23.7	276687	4	US-09-949-016-13840	Sequence 13840, A
37	27	23.5	21234	3	US-09-810-671-3	Sequence 3, Appli
38	27	23.5	21234	4	US-10-109-854-3	Sequence 3, Appli
39	27	23.5	21234	4	US-10-339-656-3	Sequence 3, Appli
c 40	27	23.5	92505	4	US-09-949-016-14018	Sequence 14018, A
41	27	23.5	130971	4	US-09-949-016-14205	Sequence 14205, A
42	27	23.5	156324	4	US-09-949-016-13749	Sequence 13749, A
c 43	27	23.5	390416	4	US-09-949-016-16923	Sequence 16923, A
c 44	26.8	23.3	601	4	US-09-949-016-44719	Sequence 44719, A
c 45	26.8	23.3	22390	4	US-09-949-016-15311	Sequence 15311, A

ALIGNMENTS

RESULT 1

US-09-831-630-8
; Sequence 8, Application US/09831630
; Patent No. 6800468
; GENERAL INFORMATION:
; APPLICANT: Clausen, Henrik
; APPLICANT: Amado, Margarita
; TITLE OF INVENTION: UDP-GALACTOSE: BETA-N-ACETYL-GLUCOSAMINE BETA 1,3
; FILE REFERENCE: 7188-157
; CURRENT APPLICATION NUMBER: US/09/831,630
; CURRENT FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 8
; LENGTH: 1011
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (79)..(1008)
US-09-831-630-8

Query Match 100.0%; Score 115; DB 4; Length 1011;
Best Local Similarity 100.0%; Pred. No. 3.2e-31;
Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CCACCTCAGCCTCCTAGCATAAACTAGACACATCCTCATCCTTTTGGAGGCTTAATCATT 60
Db 1 CCACCTCAGCCTCCTAGCATAAACTAGACACATCCTCATCCTTTTGGAGGCTTAATCATT 60
Qy 61 GGATTTTGTTCCTTTCAGATGGCTTTCCGAGAGATGAGATTGATATATCTGCC 115
Db 61 GGATTTTGTTCCTTTCAGATGGCTTTCCGAGAGATGAGATTGATATATCTGCC 115

RESULT 2

US-09-270-767-2980
; Sequence 2980, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2980
; LENGTH: 327

```
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-2980

Query Match      26.3%; Score 30.2; DB 4; Length 327;
Best Local Similarity 56.6%; Pred. No. 0.49;
Matches 56; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

Qy 2 CACCTCAGCTCCCTAGCATAAACATAGACACATCTCTCATGCTTTTGAGGTCTAATCATTTG 61
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 163 CACCTAATCTAGTAGCATAAATGTAAGACAGACAAATGCAGATAATGAATTCCTCATTT 222
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 62 GATTTTGTTCTTTTCAGATGGCTTTCCCGAAGATGAGAT 100
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 223 CATTTAATTCAAAGCAAAACCGATTCCCAAGAGGGCAAT 261
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 3
US-09-270-767-18262
; Sequence 18262, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18262
; LENGTH: 327
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-18262

Query Match      26.3%; Score 30.2; DB 4; Length 327;
Best Local Similarity 56.6%; Pred. No. 0.49;
Matches 56; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

Qy 2 CACCTCAGCTCCCTAGCATAAACATAGACACATCTCTCATGCTTTTGAGGTCTAATCATTTG 61
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 163 CACCTAATCTAGTAGCATAAATGTAAGACAGACAAATGCAGATAATGAATTCCTCATTT 222
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 62 GATTTTGTTCTTTTCAGATGGCTTTCCCGAAGATGAGAT 100
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 223 CATTTAATTCAAAGCAAAACCGATTCCCAAGAGGGCAAT 261
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 4
US-09-762-311-2/C
; Sequence 2, Application US/09762311
; Patent No. 6825004
; GENERAL INFORMATION:
; APPLICANT: BLUMENFELD, Marta
; APPLICANT: BOUGUELERET, Lydie
; APPLICANT: CHUMAKOV, Ilya
; TITLE OF INVENTION: Nucleic Acids Encoding Human TBC-1 Protein And Polymorphic Markers
; FILE REFERENCE: 46.US2.PCT
; CURRENT APPLICATION NUMBER: US/09/762,311
; CURRENT FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/095,653
; PRIOR FILING DATE: 1998-08-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent.pm
; SEQ ID NO 2
; LENGTH: 99960
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon
; LOCATION: 4661..4789
; OTHER INFORMATION: exon A
; NAME/KEY: exon
; LOCATION: 6116..6202
; OTHER INFORMATION: exon B
; NAME/KEY: exon
; LOCATION: 9919..10199
; OTHER INFORMATION: exon C
; NAME/KEY: exon
; LOCATION: 14521..14660
; OTHER INFORMATION: exon D
; NAME/KEY: exon
; LOCATION: 50257..50442
; OTHER INFORMATION: exon E
; NAME/KEY: exon
; LOCATION: 56256..56417
; OTHER INFORMATION: exon F
; NAME/KEY: exon
; LOCATION: 63326..63484
; OTHER INFORMATION: exon G
; NAME/KEY: exon
; LOCATION: 76036..76280
; OTHER INFORMATION: exon H
; NAME/KEY: exon
; LOCATION: 78364..78523
; OTHER INFORMATION: exon I
; NAME/KEY: exon
; LOCATION: 85295..85464
; OTHER INFORMATION: exon J
; NAME/KEY: exon
; LOCATION: 93417..93590
; OTHER INFORMATION: exon K
; NAME/KEY: exon
; LOCATION: 97476..97960
; OTHER INFORMATION: exon L
; NAME/KEY: misc feature
; LOCATION: 97961..99960
; OTHER INFORMATION: 3' regulatory region
; NAME/KEY: allele
; LOCATION: 1443
; OTHER INFORMATION: 99-20508-456 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 5247
; OTHER INFORMATION: 99-20469-213 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 6223
; OTHER INFORMATION: 5-254-227 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 14723
; OTHER INFORMATION: 5-257-353 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 19186
; OTHER INFORMATION: 99-20511-32 : polymorphic base C or T
; NAME/KEY: allele
; LOCATION: 18997
; OTHER INFORMATION: 99-20511-221 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 19891
; OTHER INFORMATION: 99-20510-115 : deletion of TCT
; NAME/KEY: allele
; LOCATION: 29617
; OTHER INFORMATION: 99-20504-90 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 42519
; OTHER INFORMATION: 99-20493-238 : polymorphic base A or C
; NAME/KEY: allele
; LOCATION: 69324
; OTHER INFORMATION: 99-20499-221 : polymorphic base A or G
; NAME/KEY: allele
; LOCATION: 69181
; OTHER INFORMATION: 99-20499-364 : polymorphic base A or T
; NAME/KEY: allele
; LOCATION: 69146
; OTHER INFORMATION: 99-20499-399 : polymorphic base A or G
; NAME/KEY: allele
```

```

: OTHER INFORMATION: 99-20473.pu
: NAME/KEY: primer_bind
: LOCATION: 76771..76790
: OTHER INFORMATION: 99-20473.rp complement
: NAME/KEY: primer_bind
: LOCATION: 78292..78309
: OTHER INFORMATION: 5-249.pu
: NAME/KEY: primer_bind
: LOCATION: 78704..78721
: OTHER INFORMATION: 5-249.rp complement
: NAME/KEY: primer_bind
: LOCATION: 81893..81912
: OTHER INFORMATION: 99-20485.pu
: NAME/KEY: primer_bind
: LOCATION: 82353..82372
: OTHER INFORMATION: 99-20485.rp complement
: NAME/KEY: primer_bind
: LOCATION: 84392..84412
: OTHER INFORMATION: 99-20481.pu
: NAME/KEY: primer_bind
: LOCATION: 84909..84929
: OTHER INFORMATION: 99-20481.rp complement
: NAME/KEY: primer_bind
: LOCATION: 89746..89765
: OTHER INFORMATION: 99-20480.rp
: NAME/KEY: primer_bind
: LOCATION: 90179..90198
: OTHER INFORMATION: 99-20480.pu complement
: NAME/KEY: primer_bind
: LOCATION: 94755..9493
: OTHER INFORMATION: 99-430-352.mis

Query Match          26.1%; Score 30; DB 4; Length 99960;
Best Local Similarity 54.5%; Pred. No. 6.4;
Matches 60; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

Qy 1 CCACCTCAGCCTCTAGCATATAAACTAGACACATCCTCATGCTTTTGGTCTTAATCATTT 60
Db 90492 CCAGCTCAACCTACCATCATCATATTAATCCAGTCATCAAGTGTAGACATCTAGTTCCA 90433

Qy 61 GGATTTTGTCTTTTCAGATGCTTTCCCGAGATGAGATTGATGTATAT 110
Db 90432 AGTTCTGTATTATAAGTGACAAGCTAGGAGATATCCTTTATACATAT 90383

RESULT 5
US-08-956-171E-2970
: Sequence 2970, Application US/08956171E
: Patent No. 6593114
: GENERAL INFORMATION:
: APPLICANT: Charles Kunsch
: Gil H. Choi
: Patrick S. Dillon
: Craig A. Rosen
: Steven C. Barash
: Michael R. Fannon
: TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
: NUMBER OF SEQUENCES: 5256
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Human Genome Sciences, Inc.
: STREET: 9410 Key West Avenue
: CITY: Rockville
: STATE: Maryland
: COUNTRY: USA
: ZIP: 20850
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
: COMPUTER: HP Vectra 486/33
: OPERATING SYSTEM: MSDOS version 6.2
: SOFTWARE: ASCII Text
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/956,171E
: FILING DATE: 20-Oct-1997

```

```
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 2970:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 359 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 2970:
US-08-956-171E-2970

Query Match 25.2%; Score 29; DB 4; Length 359;
Best Local Similarity 54.9%; Pred. No. 1.4;
Matches 56; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 10 CCTCTAGCATAAACTAGACACATCCTCATGCTTTTGAGGTCTAATCATTTGGATTGTGT 69
DB 231 CATCATACCATATAACTTTTATCATATAATCATTTTCAGCGAATTTAGTTGATGTTTTC 290
QY 70 TCCTTTTCAGATGGCTTCCCGAAGATGAGATTGATGATATC 111
DB 291 TTGATTAAATCTTTTCGGCGANCCCTCAGCTTGATGTTTTC 332

RESULT 7
US-09-120-653D-4
; Sequence 4, Application US/09120653D
; Patent No. 6365727
; GENERAL INFORMATION:
; APPLICANT: YOON, JI-WON
; APPLICANT: JUN, HEE-SOOK
; APPLICANT: PARK, HAE-JOON
; APPLICANT: AHN, JONG-SEONG
; APPLICANT: HA, YOUNG-JU
; APPLICANT: CHUNG, SOO-IL
; TITLE OF INVENTION: DNA and peptides of a diabetes-specific endogenous retrovirus
; FILE REFERENCE: 98-338
; CURRENT APPLICATION NUMBER: US/09/120,653D
; CURRENT FILING DATE: 1998-07-22
; PRIOR APPLICATION NUMBER: KR 98-10108
; PRIOR FILING DATE: 1998-03-24
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: KopatentIn 1.71
; SEQ ID NO 4
; LENGTH: 2330
; TYPE: DNA
; ORGANISM: DIABETES-SPECIFIC RETROVIRUS ENDOGENOUS ERV-9
US-09-120-653D-4

Query Match 25.0%; Score 28.8; DB 3; Length 2330;
Best Local Similarity 62.5%; Pred. No. 3.5;
Matches 45; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 16 AGCATAAACTAGACACATCCTCATGCTTTTGAGGTCTAATCATTTGGATTGTGTCTTT 75
DB 179 AGCTATAACCAAGTATTTCTCCACCTCTCAGTTGTAATTGGGAGACTTTGCTCTTTT 238
QY 76 CAGATGGCTTTC 87
DB 239 CACATGCCTTTC 250

RESULT 8
US-09-120-653D-1
; Sequence 1, Application US/09120653D
; Patent No. 6365727
; GENERAL INFORMATION:
; APPLICANT: YOON, JI-WON
; APPLICANT: JUN, HEE-SOOK
; APPLICANT: PARK, HAE-JOON
; APPLICANT: AHN, JONG-SEONG
; APPLICANT: HA, YOUNG-JU
; APPLICANT: CHUNG, SOO-IL
; TITLE OF INVENTION: DNA and peptides of a diabetes-specific endogenous retrovirus
; FILE REFERENCE: 98-338
; CURRENT APPLICATION NUMBER: US/09/120,653D
; CURRENT FILING DATE: 1998-07-22
; PRIOR APPLICATION NUMBER: KR 98-10108
; PRIOR FILING DATE: 1998-03-24
; NUMBER OF SEQ ID NOS: 33
```

```
; SOFTWARE: KopatentIn 1.71
; SEQ ID NO 1
; LENGTH: 3910
; TYPE: DNA
; ORGANISM: DIABETES-SPECIFIC ENDOGENOUS RETROVIRUS ERV-9
US-09-120-653D-1

Query Match      25.0%; Score 28.4; DB 3; Length 3910;
Best Local Similarity 62.5%; Pred. No. 4.4;
Matches 45; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

Qy 16 ACCATAAACTAGACACATCCTCATGCTTTTGAGGTCTAAATCATTTGGATTTTTCCTTT 75
Db 575 AGCCTATAACAGGTAATTTCTCCACCTCCTCAGTTGTAATTTGGAGACTTTTGCTCTTTT 634

Qy 76 CAGATGGCTTTC 87
Db 635 CACATGCCCTTTC 646

RESULT 9
US-09-949-016-16928
; Sequence 16928, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16928
; LENGTH: 524032
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(524032)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16928

Query Match      24.7%; Score 28.4; DB 4; Length 524032;
Best Local Similarity 56.4%; Pred. No. 47;
Matches 53; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

Qy 15 TAGCATAAACTAGACACATCCTCATGCTTTTGAGGTCTAAATCATTTGGATTTTTCCTTT 74
Db 270680 TACTAAATAAGTAGTAACAAAAAATGATGTTTATTTTCATTATATTATTTGATTTT 270739

Qy 75 TCAGATGCTTTCCGAGAGATGAGATTGATGTAT 108
Db 270740 CTAATGGGTGCTCTAGGAAGAAATGATACAT 270773

RESULT 10
US-09-949-016-16929
; Sequence 16929, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
```

```
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16929
; LENGTH: 524032
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(524032)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16929

Query Match      24.7%; Score 28.4; DB 4; Length 524032;
Best Local Similarity 56.4%; Pred. No. 47;
Matches 53; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

Qy 15 TAGCATAAACTAGACACATCCTCATGCTTTTGAGGTCTAAATCATTTGGATTTTTCCTTT 74
Db 270680 TACTAAATAAGTAGTAACAAAAAATGATGTTTATTTTCATTATATTATTTGATTTT 270739

Qy 75 TCAGATGCTTTCCGAGAGATGAGATTGATGTAT 108
Db 270740 CTAATGGGTGCTCTAGGAAGAAATGATACAT 270773

RESULT 11
US-09-949-016-16930
; Sequence 16930, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16930
; LENGTH: 524032
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(524032)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16930

Query Match      24.7%; Score 28.4; DB 4; Length 524032;
Best Local Similarity 56.4%; Pred. No. 47;
Matches 53; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

Qy 15 TAGCATAAACTAGACACATCCTCATGCTTTTGAGGTCTAAATCATTTGGATTTTTCCTTT 74
Db 270680 TACTAAATAAGTAGTAACAAAAAATGATGTTTATTTTCATTATATTATTTGATTTT 270739

Qy 75 TCAGATGCTTTCCGAGAGATGAGATTGATGTAT 108
Db 270740 CTAATGGGTGCTCTAGGAAGAAATGATACAT 270773
```



```
RESULT 12
US-09-949-016-16931
; Sequence 16931, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16931
; LENGTH: 524032
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc feature
; LOCATION: (1)-(524032)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16931

Query Match      24.7%; Score 28.4; DB 4; Length 524032;
Best Local Similarity 56.4%; Pred. No. 47;
Matches 53; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

Qy      15 TAGCATAAACTAGACACATCCTCATGCTTTTGGAGTCTAATCATGATTTTGTCTT 74
Db      270680 TACTAAATAAGTAGTAACAAAAAATGATGTTTATTTTCATTCATTATATTGTGATTTT 270739

Qy      75 TCAGATGCTTTCCGAGAGATGAGATTGATGAT 108
Db      270740 CTAATGGGTGCTCTAGGAGGAAAATGATACAT 270773

RESULT 13
US-09-949-016-14340
; Sequence 14340, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14340
; LENGTH: 529885
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc feature
; LOCATION: (1)-(529885)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14340

Query Match      24.7%; Score 28.4; DB 4; Length 529885;
Best Local Similarity 56.4%; Pred. No. 48;
Matches 53; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

Qy      15 TAGCATAAACTAGACACATCCTCATGCTTTTGGAGTCTAATCATGATTTTGTCTT 74
Db      270680 TACTAAATAAGTAGTAACAAAAAATGATGTTTATTTTCATTCATTATATTGTGATTTT 270739

Qy      75 TCAGATGCTTTCCGAGAGATGAGATTGATGAT 108
Db      270740 CTAATGGGTGCTCTAGGAGGAAAATGATACAT 270773

RESULT 14
US-09-949-016-14341
; Sequence 14341, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14341
; LENGTH: 529885
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc feature
; LOCATION: (1)-(529885)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14341

Query Match      24.7%; Score 28.4; DB 4; Length 529885;
Best Local Similarity 56.4%; Pred. No. 48;
Matches 53; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

Qy      15 TAGCATAAACTAGACACATCCTCATGCTTTTGGAGTCTAATCATGATTTTGTCTT 74
Db      270680 TACTAAATAAGTAGTAACAAAAAATGATGTTTATTTTCATTCATTATATTGTGATTTT 270739

Qy      75 TCAGATGCTTTCCGAGAGATGAGATTGATGAT 108
Db      270740 CTAATGGGTGCTCTAGGAGGAAAATGATACAT 270773

RESULT 15
US-09-949-016-14342
; Sequence 14342, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14342
; LENGTH: 529885
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc feature
; LOCATION: (1)-(529885)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14342

Query Match      24.7%; Score 28.4; DB 4; Length 529885;
Best Local Similarity 56.4%; Pred. No. 48;
Matches 53; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

Qy      15 TAGCATAAACTAGACACATCCTCATGCTTTTGGAGTCTAATCATGATTTTGTCTT 74
Db      270680 TACTAAATAAGTAGTAACAAAAAATGATGTTTATTTTCATTCATTATATTGTGATTTT 270739

Qy      75 TCAGATGCTTTCCGAGAGATGAGATTGATGAT 108
Db      270740 CTAATGGGTGCTCTAGGAGGAAAATGATACAT 270773
```

```

; LENGTH: 529885
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(529885)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14342

Query Match      24.7%; Score 28.4; DB 4; Length 529885;
Best Local Similarity 56.4%; Pred. No. 48;
Matches 53; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

Qy      15 TAGCATAAACTAGACACATCCTCATGCTTTTGAGGCTCTAATCATTTGGATTTCCTT 74
Db      270680 TACTAAATAAGTAGTACAAAAAATGATGTTTATTTTCATTTCATTATATTGTGATTTT 270739

Qy      75 TCAGATGGCTTTCCGAGAGATGAGATTGATGTAT 108
Db      270740 CTAAATGGGTGTCCTTAGGAGGAAAAATGATACAT 270773

```

Search completed: September 22, 2005, 17:02:04
Job time : 31.9415 secs

This Page Blank (uspro)



GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2005, 14:50:12 ; Search time 298.529 Seconds
(without alignments)
2575.655 Million cell updates/sec

Title: US-10-777-828-8_COPY_1_115

Perfect score: 115

Sequence: 1 ccactcagcctctagcat.....gagattgatgtatctgcc 115

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 7400704 seqs, 3343079526 residues

Total number of hits satisfying chosen parameters: 14801408

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
- 20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
- 21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
- 22: /cgn2_6/ptodata/1/pubpna/US10J_NEW_PUB.seq.*
- 23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
- 24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	115	100.0	1011	19	US-10-777-828-8
2	113.4	98.6	596	16	Sequence 13121, A
3	33	28.7	78025	14	US-10-020-141-9
4	32	27.8	495269	17	US-10-398-221-8
5	32	27.8	3011208	17	US-10-398-221-2058
6	30.8	26.8	3796	19	US-10-433-793-158
7	30.6	26.6	2000	9	US-09-938-842A-4201

c 9	30.6	26.6	2000	11	US-09-938-842A-4201	Sequence 4201, Ap
c 10	30.4	26.4	2000	9	US-09-938-842A-3143	Sequence 3143, Ap
c 11	30.4	26.4	2000	11	US-09-938-842A-3143	Sequence 3143, Ap
c 12	30.2	26.3	335199	21	US-10-496-011-4	Sequence 4, Appli
c 13	29.8	26.3	260027	13	US-10-087-192-298	Sequence 298, App
c 14	29.8	25.9	389	17	US-10-242-535A-20929	Sequence 20929, A
c 15	29.8	25.9	389	18	US-10-085-783A-20929	Sequence 20929, A
c 16	29.8	25.9	682	15	US-10-106-698-604	Sequence 604, App
c 17	29.8	25.9	4506	18	US-10-467-042-27	Sequence 27, Appli
c 18	29.4	25.6	568	22	US-11-046-868-27	Sequence 27, Appli
c 19	29.4	25.6	600	22	US-10-972-079-94291	Sequence 94291, A
c 20	29.4	25.6	600	22	US-10-972-079-94290	Sequence 94290, A
c 21	29.4	25.6	1286	9	US-09-755-016-7	Sequence 7, Appli
c 22	29.4	25.6	1286	24	US-11-049-613-7	Sequence 7, Appli
c 23	29.4	25.6	1443	16	US-10-114-153-81	Sequence 81, Appli
c 24	29.2	25.4	487	13	US-10-450-826-46	Sequence 46, Appli
c 25	29.2	25.4	487	13	US-10-027-632-276926	Sequence 276926,
c 26	29.2	25.4	487	13	US-10-027-632-276927	Sequence 276927,
c 27	29.2	25.4	487	17	US-10-027-632-276926	Sequence 276926,
c 28	29.2	25.4	799	20	US-10-027-632-276927	Sequence 276927,
c 29	29.2	25.4	5683	15	US-10-425-115-174734	Sequence 174734,
c 30	29.2	25.4	5683	15	US-10-311-455-1780	Sequence 1780, Ap
c 31	29.2	25.4	71843	22	US-10-240-485-144	Sequence 144, App
c 32	29	25.2	359	8	US-10-488-292-3	Sequence 3, Appli
c 33	29	25.2	359	18	US-08-781-986A-2970	Sequence 2970, Ap
c 34	29	25.2	1334	18	US-10-329-624-2970	Sequence 2970, Ap
c 35	29	25.2	3251	13	US-10-424-599-138167	Sequence 138167,
c 36	29	25.2	3251	13	US-10-027-632-115828	Sequence 115828,
c 37	29	25.2	3251	13	US-10-027-632-115829	Sequence 115829,
c 38	29	25.2	3251	13	US-10-027-632-115830	Sequence 115830,
c 39	29	25.2	3251	17	US-10-027-632-115828	Sequence 115828,
c 40	29	25.2	3251	17	US-10-027-632-115829	Sequence 115829,
c 41	28.8	25.0	602	13	US-10-027-632-247342	Sequence 247342,
c 42	28.8	25.0	602	13	US-10-027-632-247343	Sequence 247343,
c 43	28.8	25.0	602	13	US-10-027-632-247344	Sequence 247344,
c 44	28.8	25.0	602	13	US-10-027-632-247345	Sequence 247345,
c 45	28.8	25.0	602	17	US-10-027-632-247342	Sequence 247342,

ALIGNMENTS

RESULT 1
US-10-777-828-8
; Sequence 8, Application US/10777828
; Publication No. US20040142425A1
; GENERAL INFORMATION:
; APPLICANT: Clausen, Henrik
; APPLICANT: Amado, Margarita
; TITLE OF INVENTION: UDP-GALACTOSE: BETA-N-ACETYL-GLUCOSAMINE BETA 1,3
; FILE REFERENCE: 7188-157
; CURRENT APPLICATION NUMBER: US/10/777,828
; CURRENT FILING DATE: 2004-02-12
; PRIOR APPLICATION NUMBER: US/09/831,630
; PRIOR FILING DATE: 2001-05-10
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 1011
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (79)..(1008)
US-10-777-828-8

Query Match 100.0%; Score 115; DB 19; Length 1011;
Best Local Similarity 100.0%; Pred. No. 1.5e-28;
Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CCACCTCAGCCTCCTAGCATAAACTAGACACATCCTCATGCTTTTGGAGTCTAATCATT 60

```
Db 1 CCACCTGAGCTCTTAGCATAAACTAGACACATCTCTGCTTTGAGGTCCTAATCATT 60
Qy 61 GGATTTTGTCTCTTCAGATGGCTTCCGGAAGATGAGATTGATGATATCTGCC 115
Db 61 GGATTTTGTCTCTTCAGATGGCTTCCGGAAGATGAGATTGATGATATCTGCC 115

RESULT 2
US-10-029-386-13121
; Sequence 13121, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: AEMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029.386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 13121
; LENGTH: 596
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR21.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.71
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.53
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.59
; OTHER INFORMATION: NT HIT: g116170760, EVALUATE 0.00e+00
; OTHER INFORMATION: EST HUMAN HIT: AJ003597.1, EVALUATE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: Q9Y2C3, EVALUATE 3.00e-59
US-10-029-386-13121

Query Match 98.6%; Score 113.4; DB 16; Length 596;
Best Local Similarity 99.1%; Pred. No. 4.3e-28;
Matches 114; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CCACCTGAGCTCTTAGCATAAACTAGACACATCTCTGCTTTGAGGTCCTAATCATT 60
Db 187 CCACCTGAGCTCTTAGCATAAACTAGACACATCTCTGCTTTGAGGTCCTAATCATT 246

Qy 61 GGATTTTGTCTCTTCAGATGGCTTCCGGAAGATGAGATTGATGATATCTGCC 115
Db 247 GGATTTTGTCTCTTCAGATGGCTTCCGGAAGATGAGATTGATGATATCTGCC 301

RESULT 3
US-10-020-141-9/c
; Sequence 9, Application US/10020141
; Publication No. US2003092013A1
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Jeanette
; APPLICANT: Ableson, Allen
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF VASCULAR DISEASE
; FILE REFERENCE: MMI-002
; CURRENT APPLICATION NUMBER: US/10/020.141
; CURRENT FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/313,097
; PRIOR FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: US 60/327,485
; PRIOR FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 78025
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-020-141-9
```

```
Query Match 28.7%; Score 33; DB 14; Length 78025;
Best Local Similarity 65.8%; Pred. No. 6.8;
Matches 48; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Qy 15 TAGCATAAACTAGACACATCTCTGCTTTTGGAGGTCCTAATCATTGATTTTCTCTTT 74
Db 5992 TAGCTGTAAACAGGTACTTCTCCACCTTCTCACTTGTAAATGGGAGACTTTTCTCTTT 5933

Qy 75 TCAGATGGCTTTC 87
Db 5932 TCACATGGCTTTC 5920

RESULT 4
US-10-398-221-8
; Sequence 8, Application US/10398221
; Publication No. US20040018514A1
; GENERAL INFORMATION:
; APPLICANT: KUNST, Frederik
; APPLICANT: GLASER, Philippe
; TITLE OF INVENTION: Listeria innocua, genome and applications
; FILE REFERENCE: 344 702 - US
; CURRENT APPLICATION NUMBER: US/10/398.221
; CURRENT FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: FR 00/12 697
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 4025
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 495269
; TYPE: DNA
; ORGANISM: Listeria innocua
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(end)
; OTHER INFORMATION: n can be any nucleotide: a, g, c or t/u
US-10-398-221-8

Query Match 27.8%; Score 32; DB 17; Length 495269;
Best Local Similarity 58.3%; Pred. No. 31;
Matches 56; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

Qy 16 AGCATAAACTAGACACATCTCTGCTTTTGGAGGTCCTAATCATTGATTTTCTCTTT 75
Db 390225 AGCATACAACGACTCATAGAGAAATGCACCTGCTTGTCTAATGTTGATTTATTGATA 390284

Qy 76 CAGATGGCTTTCCGGAAGATGAGATTGATGATATC 111
Db 390285 GCCTGGTATGCGTGAATTCGGGATTCGTTTAAATC 390320

RESULT 5
US-10-398-221-2058/c
; Sequence 2058, Application US/10398221
; Publication No. US20040018514A1
; GENERAL INFORMATION:
; APPLICANT: KUNST, Frederik
; APPLICANT: GLASER, Philippe
; TITLE OF INVENTION: Listeria innocua, genome and applications
; FILE REFERENCE: 344 702 - US
; CURRENT APPLICATION NUMBER: US/10/398.221
; CURRENT FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: FR 00/12 697
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 4025
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2058
; LENGTH: 3011208
; TYPE: DNA
```

ORGANISM: Listeria innocua
US-09-938-842A-4201

Query Match 27.8%; Score 32; DB 17; Length 3011208;
Best Local Similarity 58.3%; Pred. No. 64;
Matches 56; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
QY 16 AGCATAAACTAGACACATCCCTGCTTTGAGGTCTAATCATTTGGATTGTTTCCTTT 75
DB 1368725 AGCATACACGACTCATGAGAAATGCACCTTCTGCTAATGTTGGATTATTATTGATA 1368666
QY 76 CAGATGGCTTCCCGAAGATGAGATTGATGATATC 111
DB 1368665 CCGCTGGTATCGTGAATCGGGATTGGTTAAATC 1368630

RESULT 6

US-10-433-793-158
; Sequence 158, Application US/10433793
; Publication No. US20040142334A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von mit Angiogenese assoziierten Krankheiten
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/433,793
; CURRENT FILING DATE: 2003-06-06
; NUMBER OF SEQ ID NOS: 212
; SEQ ID NO 158
; LENGTH: 3796
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-433-793-158

Query Match 26.8%; Score 30.8; DB 19; Length 3796;
Best Local Similarity 61.0%; Pred. No. 11;
Matches 50; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
QY 31 ACATCCTCATGCTTTTGAGGTCTAATCATTTGTTCTTCATGATGCTTTCCG 90
DB 3311 AAATGTAAGTATATGTTATCTATCTAGGATTTTTTTTTTAAAGATTTTAAATG 3370
QY 91 AAGATGAGATTGATATATCT 112
DB 3371 GAGGTGGATTGATTTAAATAT 3392

RESULT 7

US-09-938-842A-4201
; Sequence 4201, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 4201
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana

US-09-938-842A-4201

Query Match 26.6%; Score 30.6; DB 9; Length 2000;
Best Local Similarity 55.0%; Pred. No. 10;
Matches 60; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
QY 2 CACCTCAGCCTCTAGCATATAAACTAGACACATCTCTCATGCTTTTGAGGTCTAATCATTTG 61
DB 1077 CACCAAAATAAATCAAGCACGGAACCTCGAAAGATGATGATGCTTTCTTGTCTAATTTATG 1136
QY 62 GATTTTGTCTTTTCAGATGCTTTCCGGAAGATGAGATTGATGATAT 110
DB 1137 TGGATTGAATCATTTGGATTGGTTACCAAAAAAGAGAGATAATTT 1185

RESULT 8

US-09-938-842A-4201
; Sequence 4201, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 4201
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-4201

Query Match 26.6%; Score 30.6; DB 11; Length 2000;
Best Local Similarity 55.0%; Pred. No. 10;
Matches 60; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
QY 2 CACCTCAGCCTCTAGCATATAAACTAGACACATCTCTCATGCTTTTGAGGTCTAATCATTTG 61
DB 1077 CACCAAAATAAATCAAGCACGGAACCTCGAAAGATGATGATGCTTTCTTGTCTAATTTATG 1136
QY 62 GATTTTGTCTTTTCAGATGCTTTCCGGAAGATGAGATTGATGATAT 110
DB 1137 TGGATTGAATCATTTGGATTGGTTACCAAAAAAGAGAGATAATTT 1185

RESULT 9

US-09-938-842A-3143/c
; Sequence 3143, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16

```
; TITLE OF INVENTION: POLYNUCLEOTIDE AND PROTEIN INVOLVED IN SYNAPTOGENESIS, VARIANTS
; TITLE OF INVENTION: THEREOF, AND THEIR THERAPEUTIC AND DIAGNOSTIC USES
; FILE REFERENCE: 253820USXPCT
; CURRENT APPLICATION NUMBER: US/10/496,011
; CURRENT FILING DATE: 2004-05-28
; PRIOR APPLICATION NUMBER: PCT/FR02/04134
; PRIOR FILING DATE: 2002-11-28
; PRIOR APPLICATION NUMBER: CA2364106
; PRIOR FILING DATE: 2001-11-30
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4
; LENGTH: 335199
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon
; LOCATION: (10156)..(10298)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (108500)..(109001)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (205869)..(205928)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (209526)..(209679)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (235028)..(235139)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (238090)..(238212)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (310597)..(310783)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (316139)..(316929)
; FEATURE:
; NAME/KEY: exon
; LOCATION: (326822)..(330136)
; US-10-496-011-4
;
; Query Match 26.4%; Score 30.4; DB 21; Length 335199;
; Best Local Similarity 63.9%; Pred. No. 93;
; Matches 46; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
;
Qy 16 AGCATAAACTAGACACATCTCATGCTTTTGGAGTCTTAATCATTTGATTTGTTCTTTT 75
Db 304789 AGCTGTAAACAGGATTTCTCCAGCTCTCACTTGTAAATGGGAGACTTTGCTCTTT 304730
Qy 76 CAGATGGCTTTC 87
Db 304729 CACATGCTTTC 304718
;
RESULT 12
US-10-087-192-298/c
; Sequence 298, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
;
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 8143
; LENGTH: 2090
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-09-938-842A-3143
;
; Query Match 26.4%; Score 30.4; DB 11; Length 2000;
; Best Local Similarity 59.1%; Pred. No. 12;
; Matches 52; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
;
Qy 19 ATAAACTAGACACATCTCATGCTTTTGGAGTCTTAATCATTTGATTTGTTCTTTT 78
Db 203 ATTTATCTGTCTAATCTCACTTTTACCTGTTTGATCTTGGGTTTGCTTCTTAT 144
Qy 79 ATGGCTTTCCCGAAGATGAGATTGATGT 106
Db 143 GTCGGCTTTGGGAGATTGGTTGATTT 116
;
RESULT 10
US-09-938-842A-3143/c
; Sequence 3143, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Krips, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 8143
; LENGTH: 2090
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-09-938-842A-3143
;
; Query Match 26.4%; Score 30.4; DB 11; Length 2000;
; Best Local Similarity 59.1%; Pred. No. 12;
; Matches 52; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
;
Qy 19 ATAAACTAGACACATCTCATGCTTTTGGAGTCTTAATCATTTGATTTGTTCTTTT 78
Db 203 ATTTATCTGTCTAATCTCACTTTTACCTGTTTGATCTTGGGTTTGCTTCTTAT 144
Qy 79 ATGGCTTTCCCGAAGATGAGATTGATGT 106
Db 143 GTCGGCTTTGGGAGATTGGTTGATTT 116
;
RESULT 11
US-10-496-011-4/c
; Sequence 4, Application US/10496011
; Publication No. US20050118588A1
; GENERAL INFORMATION:
; APPLICANT: BOURGERON, THOMAS
; APPLICANT: JAMAIN, STEPHANE
; APPLICANT: QUACH, HELENE
; APPLICANT: BETANCUR, CATALINA
; APPLICANT: LEBOYER, MARION
; APPLICANT: GILLBERG, CHRISTOPHER
```


; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 298
; LENGTH: 260027
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(260027)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-298

Query Match 26.3%; Score 30.2; DB 13; Length 260027;
Best Local Similarity 62.7%; Pred. No. 99;
Matches 47; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
QY 16 AGCATAAAGTAGACACATCCTCATGCTTTTGAGGCTTAATCATGATGATGTTGTTCCCTTT 75
Db 181901 AGCCTGTAAACCAAGGTATTCTCCACCTCCTGAGTTGTAATGGAGAGACTTGTCTCTTT 181842
QY 76 CAGATGGCTTTCCCG 90
Db 181841 CACATGCCTTCTCG 181827

RESULT 13
US-10-242-535A-20929
; Sequence 20929, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liaw, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/163,280
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20929
; LENGTH: 389
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (91)..(91)
; OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-20929

Query Match 25.9%; Score 29.8; DB 17; Length 389;
Best Local Similarity 59.8%; Pred. No. 9.8;
Matches 49; Conservative 0; Mismatches 33; Indels 0; Gaps 0;
QY 34 TCCTCATGCTTTTGAGGCTTAATCATGATGATGTTGTTCCCTTTCAGATGGCTTTCCCGAAG 93
Db 250 TACTGAAGTTTAAAAAATTAATCATATTTCTCTGTTCACTTTACATGTTTTTCTGGTGG 309
QY 94 ATGAGATTGATGTATATCTGCC 115
Db 310 NGTAAATGAAGTATGCTGCC 331

US-10-242-535A-20929
; Sequence 20929, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liaw, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/163,280
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20929
; LENGTH: 389
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (91)..(91)
; OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-20929

RESULT 14
US-10-085-783A-20929
; Sequence 20929, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liaw, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20929
; LENGTH: 389
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (91)..(91)
; OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-20929

Query Match 25.9%; Score 29.8; DB 18; Length 389;
Best Local Similarity 59.8%; Pred. No. 9.8;
Matches 49; Conservative 0; Mismatches 33; Indels 0; Gaps 0;
QY 34 TCCTCATGCTTTTGAGGCTTAATCATGATGATGTTGTTCCCTTTCAGATGGCTTTCCCGAAG 93
Db 250 TACTGAAGTTTAAAAAATTAATCATATTTCTCTGTTCACTTTACATGTTTTTCTGGTGG 309
QY 94 ATGAGATTGATGTATATCTGCC 115
Db 310 NGTAAATGAAGTATGCTGCC 331

RESULT 15
US-10-106-698-604
; Sequence 604, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 604
; LENGTH: 682
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (673)..(673)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-604

Query Match 25.9%; Score 29.8; DB 15; Length 682;
 Best Local Similarity 60.5%; Pred. No. 12;
 Matches 49; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
 Qy 28 GACACATCCTCATGCTTTTGAGGTCTAATCATGTGGAATTTTGTCTTTTCAGATGGCTTTC 87
 Db 456 GAGCTCCCTCATGTCCTTTAAGGTAGCATCATTTGGATGTGGCTTTTGGATTTT 515
 Qy 88 CCGAAGATGAGATTGATGTAT 108
 Db 416 CTGAACACAGCTAATGTTGTGT 536

Search completed: September 22, 2005, 22:20:24
 Job time : 303.529 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2005, 07:38:28 ; Search time 2670.69 Seconds
(without alignments)
10595.706 Million cell updates/sec

Title: US-10-777-828-8_COPY_428_1011

Perfect score: 584

Sequence: 1 gggacattatccgaagat.....aagattgcgcctgtctga 584

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_on.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	584	100.0	933	9 AF145784	Homo sapi
2	584	100.0	1011	6 B0228354	UDP-galac
3	584	100.0	2762	9 AB020337	Homo sapi
4	584	100.0	2775	6 E38419	Novel poly
5	584	100.0	2921	9 A1372061	Homo sapi
6	584	100.0	10562	6 E38420	Novel poly
7	584	100.0	170121	9 AF064860	Homo sapi
8	584	100.0	340000	9 HS21C080	Homo sapi
9	583	99.8	933	9 HSA6078	Homo sapi
10	582.4	99.7	2494	6 CQ731786	Sequence
11	569.6	97.5	192219	9 RF43002119	Pan trogl
12	553	94.7	1576	9 AB041416	Homo sapi
13	541.2	92.7	1570	9 AB041415	Pan panis
14	532.6	91.2	1360	9 AB041412	Gorilla g
15	531	90.9	1565	9 AB041413	Homo sapi
16	530.8	90.9	937	9 AY231145	Homo sapi
17	530.2	90.8	1566	9 AB041414	Macaca mu
18	530	90.8	1579	9 AB041417	Pongo pyg
19	387.2	66.3	170108	2 AC150794	Bos tauru

20	368	63.0	927	10	AF254738	Mus muscu	
21	368	63.0	4933	10	BC057887	Mus muscu	
22	368	63.0	5069	10	BC051669	Mus muscu	
23	368	63.0	196900	2	AC020851	Mus muscu	
24	363.2	62.2	149964	2	AC120145	Mus muscu	
C	25	363.2	62.2	186956	2	AC120346	Mus muscu
C	26	321.8	55.1	65400	2	AC109264	Mus muscu
27	213.8	36.6	97702	2	AC151319	Xenopus t	
C	28	159.8	27.4	226720	2	BX936311	Danio rer
29	158.2	27.1	2317	5	BC066477	Danio rer	
30	132.6	22.7	1950	5	BC082704	Xenopus l	
31	124.6	21.3	1474	5	BC075347	Xenopus t	
32	124.4	21.3	917	10	AB039137	Mus muscu	
33	124.4	21.3	917	10	AB039140	Mus muscu	
34	124.4	21.3	917	10	AB039141	Mus muscu	
35	124.4	21.3	917	10	AB039142	Mus muscu	
C	36	123.4	21.1	110000	2	AC096079_1	Continuation (2 of
C	37	123.4	21.1	231235	2	AC136678	Rattus no
C	38	123.4	21.1	231471	2	AC137010	Rattus no
39	123.4	21.1	244843	2	AC107162	Rattus no	
40	123.4	21.1	264119	2	AC134162	Rattus no	
41	123	21.1	2049	5	BC080111	Xenopus l	
42	122.8	21.0	917	10	AB039134	Mus muscu	
43	122.8	21.0	917	10	AB039135	Mus muscu	
44	122.8	21.0	917	10	AB039136	Mus muscu	
45	122.8	21.0	917	10	AB039138	Mus muscu	

ALIGNMENTS

RESULT 1
AF145784
LOCUS AF145784 933 bp DNA linear PRI 14-NOV-1999
DEFINITION Homo sapiens beta1.3 galactosyltransferase-V (B3GALT5) gene,
complete cds.

ACCESSION AF145784

VERSION AF145784.1 GI:6409192

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 933)

Zhou,D., Berger,E.G. and Hennet,T.

Molecular cloning of a human UDP-galactose-4-epimerase gene encoding an O-linked

beta1,3 galactosyltransferase gene encoding an O-linked

core3-elongation enzyme

JOURNAL Eur. J. Biochem. 263 (2), 571-576 (1999)

MEDLINE 99337698

PUBMED 10406968

REFERENCE 2 (bases 1 to 933)

Zhou,D. and Hennet,T.

Direct Submission

Submitted (26-APR-1999) Physiology, University of Zurich,

Winterthurerstrasse 190, Zurich 8057, Switzerland

Location/Qualifiers

1..933

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="21"

/map="21q22"

1..933

/gene="B3GALT5"

/note="glycosyltransferase"

/codon_start=1

/product="beta1,3 galactosyltransferase-V"

/protein_id="AA07880.1"

/db_xref="GI:6409193"

/translation="MAFFQRLMYICLLVLGALCLYFSWYSLNPFKEQSFYVKDGNF"

```

LKDPTDRCOTPPPLVLLVTSKHLAERMAIROTWMKVMKGKOLKTFLLGTSS
AAETKVDQSDRGDIIQDPLDYVNLTLKTMWGIEMWHRPCPOAFVWKTDSDMF
INVDYIELLKQNRHTFTFTFLKNEPFIQRFKSFVKSSEYFWDRIYPPFCSGTG
YVSGDVASQYNNVSKSVPIYKLDVDFVGLCLERLIRLBELHSHQPTFFPGGURFSVC
LFRIRIVAFIKPRTLLDYQWALENSRGEDCPVP"

ORIGIN
Query Match      100.0%; Score 584; DB 9; Length 933;
Best Local Similarity 100.0%; Pred. No. 6.6e-151;
Matches 584; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGACATTATCCAGAAGGATTTCCTAGACGCTCTATTACAATCTGACCCCTGAAGACCATGA 60
Db 350 GGGACATTATCCAGAAGGATTTCCTAGACGCTCTATTACAATCTGACCCCTGAAGACCATGA 409

Qy 61 TGGGCATAGAAATGGGTCCATCGCTTTTGTCTCAGGGGGCGGTTTGTGATGAAAAACAGACT 120
Db 410 TGGGCATAGAAATGGGTCCATCGCTTTTGTCTCAGGGGGCGGTTTGTGATGAAAAACAGACT 469

Qy 121 CAGACATGTTTCATCAATGTTGACTATCTGACTGAACTGCTTCTGAGAAAAACAGAACAA 180
Db 470 CAGACATGTTTCATCAATGTTGACTATCTGACTGAACTGCTTCTGAGAAAAACAGAACAA 529

Qy 181 CCAGGTTTTTCACTGGCTTCTTGAAACTCAATGAGTTTCCCATCAGGCAGGCATTTCAGCA 240
Db 530 CCAGGTTTTTCACTGGCTTCTTGAAACTCAATGAGTTTCCCATCAGGCAGGCATTTCAGCA 589

Qy 241 AGTGGTTTGTTCAGTAAATCTGAATATCCGTGGGACAGGTACCCACCAATTCGTCCGGCA 300
Db 590 AGTGGTTTGTTCAGTAAATCTGAATATCCGTGGGACAGGTACCCACCAATTCGTCCGGCA 649

Qy 301 CCGGCTACGTTTCTGGGACGTTGGGAGCTGAGGTGACAAATGTCTCAGAGGCGTCC 360
Db 650 CCGGCTACGTTTCTGGGACGTTGGGAGCTGAGGTGACAAATGTCTCAGAGGCGTCC 709

Qy 361 CATACATTAAACTGGAAGACGTTGTTGGGGCTCTGCCTCGAAAGGCTGAAACATCAGAT 420
Db 710 CATACATTAAACTGGAAGACGTTGTTGGGGCTCTGCCTCGAAAGGCTGAAACATCAGAT 769

Qy 421 TGGAGGAGCTCCACTCCAGCGGACCTTTTTCAGGGGGCTTACGCTTTCCTGATGCCC 480
Db 770 TGGAGGAGCTCCACTCCAGCGGACCTTTTTCAGGGGGCTTACGCTTTCCTGATGCCC 829

Qy 481 TCTTCAGAGAGATCGTGGCCTGCCACTTCATCAAGCCTCGGACTCTCTTGACACTAGGC 540
Db 830 TCTTCAGAGAGATCGTGGCCTGCCACTTCATCAAGCCTCGGACTCTCTTGACACTAGGC 889

Qy 541 AGGCTCTAGAGAAATCCCGGGGGGAAGATTGTCCGCCCTGTCTGA 584
Db 890 AGGCTCTAGAGAAATCCCGGGGGGAAGATTGTCCGCCCTGTCTGA 933

RESULT 2
BD228354 1011 bp DNA linear PAT 17-JUL-2003
LOCUS
DEFINITION
UDP-galactose: beta-N-acetyl-glucosamine beta 1,3
galactosyltransferases, beta 3 Gal-T5.
BD228354
BD228354.1 GI:33038124
VERSION
JP 2002530071-A/8.
KEYWORDS
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1011)
Clausen,H. and Anado,M.
UDP-galactose: beta-N-acetyl-glucosamine beta 1,3
galactosyltransferases, beta 3 Gal-T5
Patent: JP 2002530071-A 8 17-SEP-2002;
HENRIK CLAUSEN
OS Homo sapiens (human)
PN JP 2002530071-A/8
PD 17-SEP-2002
```

```

PF 11-NOV-1999 JP 2000582542
PR 13-NOV-1998 DK PA 199801483
PI HENRIK CLAUSEN, MARGARIDA AMADO
PC C12N15/09,C12N1/15,C12N1/21,C12N5/10,C12N9/10,C12N15/
PC 00,C12N5/00
CC UDP-galactose: beta-N-acetyl-glucosamine beta 1,3 CC
galactosyltransferases,
CC beta 3 Gal-T5 Location/Qualifiers
FH Key Location/Qualifiers
FT CDS (79)..(1008).

FEATURES
source
1..1011
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match      100.0%; Score 584; DB 6; Length 1011;
Best Local Similarity 100.0%; Pred. No. 6.6e-151;
Matches 584; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGACATTATCCAGAAGGATTTCCTAGACGCTCTATTACAATCTGACCCCTGAAGACCATGA 60
Db 428 GGGACATTATCCAGAAGGATTTCCTAGACGCTCTATTACAATCTGACCCCTGAAGACCATGA 487

Qy 61 TGGGCATAGAAATGGGTCCATCGCTTTTGTCTCAGGGGGCGGTTTGTGATGAAAAACAGACT 120
Db 488 TGGGCATAGAAATGGGTCCATCGCTTTTGTCTCAGGGGGCGGTTTGTGATGAAAAACAGACT 547

Qy 121 CAGACATGTTTCATCAATGTTGACTATCTGACTGAACTGCTTCTGAGAAAAACAGAACAA 180
Db 548 CAGACATGTTTCATCAATGTTGACTATCTGACTGAACTGCTTCTGAGAAAAACAGAACAA 607

Qy 181 CCAGGTTTTTCACTGGCTTCTTGAAACTCAATGAGTTTCCCATCAGGCAGGCATTTCAGCA 240
Db 608 CCAGGTTTTTCACTGGCTTCTTGAAACTCAATGAGTTTCCCATCAGGCAGGCATTTCAGCA 667

Qy 241 AGTGGTTTGTTCAGTAAATCTGAATATCCGTGGGACAGGTACCCACCAATTCGTCCGGCA 300
Db 668 AGTGGTTTGTTCAGTAAATCTGAATATCCGTGGGACAGGTACCCACCAATTCGTCCGGCA 727

Qy 301 CCGGCTACGTTTCTGGGACGTTGGGAGCTGAGGTGACAAATGTCTCAGAGGCGTCC 360
Db 728 CCGGCTACGTTTCTGGGACGTTGGGAGCTGAGGTGACAAATGTCTCAGAGGCGTCC 787

Qy 361 CATACATTAAACTGGAAGACGTTGTTGGGGCTCTGCCTCGAAAGGCTGAAACATCAGAT 420
Db 788 CATACATTAAACTGGAAGACGTTGTTGGGGCTCTGCCTCGAAAGGCTGAAACATCAGAT 847

Qy 421 TGGAGGAGCTCCACTCCAGCGGACCTTTTTCAGGGGGCTTACGCTTTCCTGATGCCC 480
Db 848 TGGAGGAGCTCCACTCCAGCGGACCTTTTTCAGGGGGCTTACGCTTTCCTGATGCCC 907

Qy 481 TCTTCAGAGAGATCGTGGCCTGCCACTTCATCAAGCCTCGGACTCTCTTGACACTAGTGC 540
Db 908 TCTTCAGAGAGATCGTGGCCTGCCACTTCATCAAGCCTCGGACTCTCTTGACACTAGTGC 967

Qy 541 AGGCTCTAGAGAAATCCCGGGGGGAAGATTGTCCGCCCTGTCTGA 584
Db 968 AGGCTCTAGAGAAATCCCGGGGGGAAGATTGTCCGCCCTGTCTGA 1011

RESULT 3
AB020337 2762 bp mRNA linear PRI 14-APR-2000
LOCUS
DEFINITION
Homo sapiens mRNA for UDP-Gal:GlcNAc beta1,3-galactosyltransferase
5, complete cds.
AB020337
AB020337.1 GI:4835502
VERSION
UDP-Gal:GlcNAc beta1,3-galactosyltransferase 5.
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (sites)
Ishiki,S., Togayachi,A., Kudo,T., Nishihara,S., Watanabe,M.,
Kubota,T., Kitajima,M., Shiraishi,N., Sasaki,K., Andoh,T. and
Narimatsu,H.

Cloning, expression, and characterization of a novel
UDP-galactose:beta-N-acetylglucosamine
betal,3-galactosyltransferase (beta3Gal-T5) responsible for
synthesis of type I chain in colorectal and pancreatic epithelia
and tumor cells derived therefrom

J. Biol. Chem. 274 (18), 12499-12507 (1999)

99230269
10212226

2 (bases 1 to 2762)
Ishiki,S., Togayachi,A. and Narimatsu,H.

Direct Submission
Submitted (20-NOV-1998) Hisashi Narimatsu, Soka University,
Institute of Life Science; 1-236, Tangi-cho, Hachioji, Tokyo
192-8577, Japan (E-mail:sishiki@po.hijnet.or.jp,
Tel:81-426-91-9466, Fax:81-426-91-9315)

Location/Qualifiers
1..2762
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
/cell_line="Colo 205"
/cell_type="Adenocarcinoma"

1..2762
/gene="betal,3-Galt 5"
1..273
/gene="betal,3-Galt 5"
/number=1
274..433
/gene="betal,3-Galt 5"
/number=3
434..2762
/gene="betal,3-Galt 5"
/number=4
434..1366
/gene="betal,3-Galt 5"
/codon_start=1
/product="UDP-Gal:GlcNAc betal,3-galactosyltransferase 5"
/protein_id="BAA77664.1"
/db_xref="GI:4835503"
/translation="MAPPKRMVYICLLVLCALCLYFSMYSINLPKQSFYKKDGNF
LKLPTDCRQTPPELVLLVTSRQLAERMAIRQWGRMVKGRKQALTFLLGTTS
AAETKVDQESQRHDIQKDFLDVYVNLTKTWGIEWHRPCQAAFAVMKTDSDMF
INVLYLLELLKQRTFFFTGFLKNEFFIRQPFKSVFSDYPPDRYPPFCSTG
YVFGSDVASQVYNSKVPYIKLEDFVGLCLERINIRLELHLSOPTFFPGGLRFSVC
LFRIVACHFIKPRLLDYQWALENSRGEDCPVP"

2762
/gene="betal,3-Galt 5"
/note="45 a nucleotides"

polyA_site

ORIGIN
Query Match 100.0%; Score 584; DB 9; Length 2762;
Best Local Similarity 100.0%; Pred. No. 6.7e-151; Indels 0; Gaps 0;
Matches 584; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGACATTATCCAGAGGATTTCCTAGACGCTCTATTACAATCTGACCTCGAGACCATGA 60
Db |||||
Qy 783 GGGACATTATCCAGAGGATTTCCTAGACGCTCTATTACAATCTGACCTCGAGACCATGA 842
Db |||||
Qy 61 TGGGCATAGAGTGGTCCATCGCTTTTCTCTAGCGCGGCTTTGTGATGAAGAAACAGACT 120
Db |||||
Qy 843 TGGGCATAGAGTGGTCCATCGCTTTTCTCTAGCGCGGCTTTGTGATGAAGAAACAGACT 902
Db |||||
Qy 121 CAGACATGTTTCAATCTGACTATCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
Db |||||
Qy 903 CAGACATGTTTCAATCTGACTATCTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 962

Qy 181 CCAGGTTTTTCACTGGCTTCTTTGAAATCAATAGTATTTCCCATCAGGAGCCATTTCAGCA 240
Db |||||
Qy 963 CCAGGTTTTTCACTGGCTTCTTTGAAATCAATAGTATTTCCCATCAGGAGCCATTTCAGCA 1022
Db |||||
Qy 241 AGTGGTTTGTGAGTAAATCTGAATATCCGTGGGACAGGTACCCACCATTTCTGCTCGGCA 300
Db |||||
Qy 1023 AGTGGTTTGTGAGTAAATCTGAATATCCGTGGGACAGGTACCCACCATTTCTGCTCGGCA 1082
Db |||||
Qy 301 CCGGCTACGTGTTTCTCGGACGCTGGGAGTCCAGTGTGTACAAATGTCTTCAAGAGCGTCC 360
Db |||||
Qy 1083 CCGGCTACGTGTTTCTCGGACGCTGGGAGTCCAGTGTGTACAAATGTCTTCAAGAGCGTCC 1142
Db |||||
Qy 361 CATACATTAAACTGGAAGACGCTGTTTGTGGGGCTCTGCTCGAAAGGCTGAACATCAGAT 420
Db |||||
Qy 1143 CATACATTAAACTGGAAGACGCTGTTTGTGGGGCTCTGCTCGAAAGGCTGAACATCAGAT 1202
Db |||||
Qy 421 TGGAGGACCTCCACTCCCGACGCGACCTTTTTCAGGGGGCTTACGCTTCTCGGTATGCC 480
Db |||||
Qy 1203 TGGAGGACCTCCACTCCCGACGCGACCTTTTTCAGGGGGCTTACGCTTCTCGGTATGCC 1262
Db |||||
Qy 481 TCTTCAGGAGGATCGTGGCTCGCCACTTTCATCAAGCCTCGGACTCTCTTGGACTACTGGC 540
Db |||||
Qy 1263 TCTTCAGGAGGATCGTGGCTCGCCACTTTCATCAAGCCTCGGACTCTCTTGGACTACTGGC 1322
Db |||||
Qy 541 AGGCTCTAGAGAAATTCGGGGGGGGAAGATTGTCCGCTGTCTGA 584
Db |||||
Qy 1323 AGGCTCTAGAGAAATTCGGGGGGGGAAGATTGTCCGCTGTCTGA 1366
Db |||||

RESULT 4
E38419
LOCUS
DEFINITION Novel polypeptide.
ACCESSION E38419
VERSION E38419.1 GI:18626993
KEYWORDS JP 2000245464-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2775)
AUTHORS Narimatsu,H., Ishiki,S., Togayachi,A. and Sasaki,K.
TITLE Novel polypeptide
JOURNAL Patent: JP 2000245464-A 1 12-SEP-2000;
KYOWA HAKKO KOGYO CO LTD
OS Homo sapiens (human)
PN JP 2000245464-A/1
PD 12-SEP-2000
PF 25-FEB-1999 JP 1999047571
PI HISASHI NARIMATSU, SOICHIRO ISSHIKI, AKIRA TOGAYACHU, PI
KATSUTOSHI SASAKI
PC C12N15/09,A01K67/027,C12N1/21,C12N5/10,C12N9/10,C12P19/00, PC
C12P21/02,
PC C12P21/08,C12Q1/68,G01N33/53//C12N1/21,C12R1/185), (C12N5/10,
PC C12R1/91),
PC (C12P21/02,C12R1/185), (C12P21/02,C12R1/91), (C12N5/00,C12N5/00,
PC (C12N5/00,C12R1/91)
CC
FH Key Location/Qualifiers
FT CDS (402)..(1331).
source
1..2775
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 584; DB 6; Length 2775;
Best Local Similarity 100.0%; Pred. No. 6.7e-151; Indels 0; Gaps 0;
Matches 584; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGACATTATCCAGAGGATTTCCTAGACGCTCTATTACAATCTGACCTCGAGACCATGA 60

TITLE Novel polypeptide
JOURNAL Patent: JP 2000245464-A 2 12-SEP-2000;
KYOWA HAKKO KOGYO CO LTD
COMMENT OS Homo sapiens (human)
PN JP 2000245464-A/2
PD 12-SEP-2000
PF 25-FEB-1999 JP 1999047571
PR
P1 HISASHI NARIMATSU, SOICHIRO ISSHIKI, AKIRA TOGAYAUCHI, PI
KATSUTOSHI SASAKI
PC C12N15/09, A01K67/027, C12N1/21, C12N5/10, C12N9/10, C12P19/00, PC
C12P21/02,
PC C12P21/08, C12Q3/68, G01N33/53/(C12N1/21, C12R1:185), (C12N5/10,
PC C12R1:91),
PC (C12P21/02, C12R1:185), (C12P21/02, C12R1:91), C12N15/00, C12N5/00,
PC (C12N5/00, C12R1:91)
CC
FH Key Location/Qualifiers
FT promoter (1)..(5000)
FT exon (5001)..(5140)
FT exon (5001)..(5273)
FT exon (5459)..(5567)
FT exon (7427)..(7586)
FT exon (8234)..(10562).
FEATURES source
1..10562
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 584; DB 6; Length 10562;
Best Local Similarity 100.0%; Pred. No. 6.9e-151;
Matches 584; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGGACATTATCCAGAGGATTCTCTAGACGCTATTACAACTGACCCCTGAAGACCATGA 60
Db 8583 GGGACATTATCCAGAGGATTCTCTAGACGCTATTACAACTGACCCCTGAAGACCATGA 8642
Qy 61 TGGGCATAGAAATGGTCCATCGCTTTCTCTCAGCGCGCTTTGTGATGAAACAGACT 120
Db 8643 TGGGCATAGAAATGGTCCATCGCTTTCTCTCAGCGCGCTTTGTGATGAAACAGACT 8702
Qy 121 CAGACATGTTCAATGCTTCACTATCTGACTATCTGACTGAACTCTTCTGAAGAAAACAGAACAA 180
Db 8703 CAGACATGTTCAATGCTTCACTATCTGACTATCTGACTGAACTCTTCTGAAGAAAACAGAACAA 8762
Qy 181 CCAGGTTTTCACTGGCTTCTTGAAACTCAATGAGTTTCCCATCAGGCGACCATTCAGCA 240
Db 8763 CCAGGTTTTCACTGGCTTCTTGAAACTCAATGAGTTTCCCATCAGGCGACCATTCAGCA 8822
Qy 241 AGTGGTTTGTGAGTAAATCTGAATATCGTGGGACAGGTACCCACCATTCCTGCTCCGGCA 300
Db 8823 AGTGGTTTGTGAGTAAATCTGAATATCGTGGGACAGGTACCCACCATTCCTGCTCCGGCA 8882
Qy 301 CCGGCTAGCTGTTTCTGGCGACGTTGGCGAGTCAAGTGTACAAATGCTTCCCAAGAGCGTCC 360
Db 8883 CCGGCTAGCTGTTTCTGGCGACGTTGGCGAGTCAAGTGTACAAATGCTTCCCAAGAGCGTCC 8942
Qy 361 CATACATTAAATCGAAGACGTTGTTGGGGCTCTGCCTCGAAAGGCTGAACATCAGAT 420
Db 8943 CATACATTAAATCGAAGACGTTGTTGGGGCTCTGCCTCGAAAGGCTGAACATCAGAT 9002
Qy 421 TGGAGGAGCTCCACTCCCGAGCGACCTTTTTCAGGGGGCTTACGCTTCTCCGATGCC 480
Db 9003 TGGAGGAGCTCCACTCCCGAGCGACCTTTTTCAGGGGGCTTACGCTTCTCCGATGCC 9062
Qy 481 TCTTCAGAGGAGTCTGGCCCTGCCACTTCATCAAGCCCTCGGACTCTCTTGGACTACTGGC 540
Db 9063 TCTTCAGAGGAGTCTGGCCCTGCCACTTCATCAAGCCCTCGGACTCTCTTGGACTACTGGC 9122
Qy 541 AGGCTCTAGAGAAATCCCGGGGGGGAAGATTGTCCGCTCTGTCTGA 584

Db 9123 AGGCTCTAGAGAAATCCCGGGGGGGAAGATTGTCCGCTGTCTGA 9166
RESULT 7
AF064860
LOCUS Homo sapiens chromosome 21 clone PAC 70124 map 21q22.3, complete
DEFINITION sequence.
ACCESSION AF064860
VERSION AF064860.2 GI:18958624
KEYWORDS HTG; HTGS DRAFT
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 170121)
AUTHORS Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T.,
Park,H.-S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.-K., Soeda,E.,
Ohki,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K.,
Polley,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R.,
Patterson,D., Reichwald,K., Rump,A., Schillhabel,M.B., Schudy,A.,
Zimmermann,W., Rosenthal,A., Kudoh,J., Kawasaki,K., Asakawa,S.,
Shintani,A., Sasaki,T., Nagamine,K., Mitsuyama,S.
Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordtsiek,G.,
Hornischer,K., Brandt,P., Sharfe,M., Schoen,O., Desario,A.,
Reichelt,J., Kauer,G., Bloecker,H., Ramser,J., Beck,A., Klages,S.,
Hennig,S., Riessekmann,L., Dagand,E., Haaf,T., Wehrmeyer,S.,
Borzym,K., Gardiner,K., Nizetic,D., Francis,F., Lehrach,H.,
Reinhardt,R. and Yaspo,M.Laure.
The DNA sequence of human chromosome 21
Nature 405 (6784), 311-319 (2000)
20289799
10830953
REFERENCE 2 (bases 1 to 170121)
AUTHORS Rump,A., Dagand,E., Hildmann,T., Nordtsiek,G., Drescher,B.,
Weber,J., Schattevoy,R., Yaspo,M.-L. and Rosenthal,A.
Direct Submission
TITLE Submitted (12-MAY-1998) Genome Analysis, Institute of Molecular
JOURNAL Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
REFERENCE 3 (bases 1 to 170121)
AUTHORS Rump,A., Dagand,E., Hildmann,T., Nordtsiek,G., Drescher,B.,
Weber,J., Schattevoy,R., Yaspo,M.-L., Rosenthal,A., Yaspo,M.-L. and
Rosenthal,A.
Direct Submission
TITLE Submitted (27-PB-2002) Genome Analysis, Institute of Molecular
JOURNAL Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
REFERENCE 4 (bases 1 to 170121)
AUTHORS Rump,A., Dagand,E., Hildmann,T., Nordtsiek,G., Drescher,B.,
Weber,J., Schattevoy,R., Yaspo,M.-L. and Rosenthal,A.
Direct Submission
TITLE Submitted (05-MAR-2002) Genome Analysis, Institute of Molecular
JOURNAL Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
COMMENT On Feb 27, 2002 this sequence version replaced gi:3171153.
FEATURES
source
1..170121
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
/clones="PAC 70124"
ORIGIN
Query Match 100.0%; Score 584; DB 9; Length 170121;
Best Local Similarity 100.0%; Pred. No. 7.2e-151;
Matches 584; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGGACATTATCCAGAGGATTCTCTAGACGCTATTACAACTGACCCCTGAAGACCATGA 60
Db 90397 GGGACATTATCCAGAGGATTCTCTAGACGCTATTACAACTGACCCCTGAAGACCATGA 90456
Qy 61 TGGGCATAGAAATGGTCCATCGCTTTTGTCTCAGCGCGCTTTGTGATGAAACAGACT 120

```
Db 90457 TGGCATAGAAATGGTCCATCGCTTTTGTCTCCTCAGCGCGGCTTTGTGATGAACACAGACT 90516
QY 121 CAGACATGTTTCATCAATGTTGACTATCTGACTGACTGACTGCTTCTGAAGAAAAACAGAACAA 180
Db 90517 CAGACATGTTTCATCAATGTTGACTATCTGACTGACTGACTGCTTCTGAAGAAAAACAGAACAA 90576
QY 181 CCAGGTTTTTCACTGGCGCTTCTTGAACATCAATGAGTTCCTCCATCAGGCGAGCCATTCAGCA 240
Db 90577 CCAGGTTTTTCACTGGCGCTTCTTGAACATCAATGAGTTCCTCCATCAGGCGAGCCATTCAGCA 90636
QY 241 AGTGGTTTGTGAGTAATCTGAATATCCGTGGGACAGGTACCCACCAATTCCTGCTCCGGCA 300
Db 90637 AGTGGTTTGTGAGTAATCTGAATATCCGTGGGACAGGTACCCACCAATTCCTGCTCCGGCA 90696
QY 301 CCGGCTACGCTGTTTTCTGGCGACGCTGGCGAGTCAAGTGTACAATGTTCTCCAAGAGCGTCC 360
Db 90697 CCGGCTACGCTGTTTTCTGGCGACGCTGGCGAGTCAAGTGTACAATGTTCTCCAAGAGCGTCC 90756
QY 361 CATACATTAAACTGGAAGACGCTGTTGTGGGGCTCTGCCTCGAAAGGCTGAACATCAGAT 420
Db 90757 CATACATTAAACTGGAAGACGCTGTTGTGGGGCTCTGCCTCGAAAGGCTGAACATCAGAT 90816
QY 421 TGGAGGAGCTCCACTCCCGAGCGACCTTTTTTTCAGGGGGCTTACGCTTCTCCGTATGCC 480
Db 90817 TGGAGGAGCTCCACTCCCGAGCGACCTTTTTTTCAGGGGGCTTACGCTTCTCCGTATGCC 90876
QY 481 TCTTCAGAGAGATGTTGGCGCTGGCACTTCATCAAGCTCGGAGCTCTCTTGGAGTACTGGC 540
Db 90877 TCTTCAGAGAGATGTTGGCGCTGGCACTTCATCAAGCTCGGAGCTCTCTTGGAGTACTGGC 90936
QY 541 AGGCTCTAGAGAAATCCCGGGGGGAAGATTCTCCGCTGCTCGA 584
Db 90937 AGGCTCTAGAGAAATCCCGGGGGGAAGATTCTCCGCTGCTCGA 90980

RESULT 8
HS21C080 340000 bp DNA linear PRI 24-MAY-2000
LOCUS Homo sapiens chromosome 21 segment HS21C080.
DEFINITION AL163280 AP001735 BA000005
VERSION AL163280.2 GI:7717369
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T.,
Park,H.S., Toyoda,A., Ishii,K., Totoki,Y., Choi,D.K., Soeda,E.,
Ohki,M., Takagi,T., Sakaki,Y., Taudien,S., Blechschmidt,K.,
Polley,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R.,
Patterson,D., Reichwald,K., Rump,A., Schillhabel,M., Schudy,A.,
Zimmermann,W., Rosenthal,A., Kudoh,J., Shibuya,K., Kawasaki,K.,
Asakawa,S., Shintani,A., Sasaki,T., Nagamine,K., Mitsuayama,S.,
Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordliek,G.,
Hornischer,K., Brandt,P., Scharfe,M., Schoen,O., Desario,A.,
Reichelt,J., Kauer,G., Bloeker,H., Ramser,J., Beck,A., Klages,S.,
Hennig,S., Riesemann,L., Dagand,E., Wehrmeyer,S., Borzym,K.,
Gardiner,K., Nizetic,D., Francis,F., Lehrach,H., Reinhardt,R. and
Yaspo,M.L.
Direct Submission
Submitted (05-MAY-2000) The Chromosome 21 Mapping and Sequencing
Consortium: * RIKEN Genomic Sciences Center, Human Genome Research
Group * Institute of Molecular Biotechnology, Genome Analysis
Keio University School of Medicine, Dept. of Molecular Biology *
GBF, Dept. of Genome Analysis * Max-Planck Institute for Molecular
Genetics (addresses see below)
The Chromosome 21 Mapping and Sequencing Consortium consists of
* RIKEN Genomic Sciences Center, Human Genome Research Group, *
Sagamihara 228-8555, Japan,
* e.mail: sakaki@gsc.riken.go.jp
* URL: http://hgp.gsc.riken.go.jp/
and
```

```
* Institute of Molecular Biotechnology, Genome Analysis, *
Butenbergrasse 11, D-07745 Jena, Germany,
* e.mail: gscj-submit@genome.imb-jena.de
* URL: http://genome.imb-jena.de/
and
* Keio University School of Medicine, Dept. of Molecular Biology, *
Tokyo 160-8582, Japan,
* e.mail: ehimizu@mb-med.keio.ac.jp
* URL: http://adenine.dmb.med.keio.ac.jp/
and
* GBF, Dept. of Genome Analysis,
* Mascheroder Weg 1, D-38124 Braunschweig, Germany, * e.mail:
info.genome@gbf.de
* URL: http://genome.gbf.de/
and
* Max-Planck Institute for Molecular Genetics,
* Ihnestrasse 73, D-14195 Berlin, Germany,
* e.mail: info-chr21@molgen.mpg.de
* URL: http://chr21.rz-berlin.mpg.de/.
FEATURES
source
1..340000
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
<1..125946
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
/clone="P160L9, 5' partial"
/clone_lib="RPC11,3-5 PAC library"
/notes="Accession No. AF121897"
73410..243533
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
/clone="P70124"
/clone_lib="RPC11,3-5 PAC library"
/notes="Accession No. AF064860"
224137..>340000
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
/clone="P206A10, 3' partial"
/clone_lib="RPC11,3-5 PAC library"
/notes="Accession No. AF121782"
286628..>340000
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.3"
/clone="BAC-291B3, 3' partial"
/clone_lib="BAC library"
/notes="Accession No. AF064857"
1088..1566
/notes="LIM4"
/rpt_family="LINE/L1"
/rpt_type=DISPERSED
2398..2495
/notes="(CTA)n"
/rpt_family="Simple_repeat"
complement(2496..2878)
/notes="THE1C"
/rpt_family="LTR/MaLR"
```



```

repeat_region      /rpt_type=DISPERSED
2879. .2931
/note="TCTA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
repeat_region      /rpt_type=DISPERSED
2932. .3004
/note="TCCA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
repeat_region      /rpt_type=DISPERSED
3005. .3074
/note="TCCA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
gene               /gene="SH3BGR"
3473. .19157
/mRNA              join(3473. .3565,12649. .12678,15331. .15460,18739. .19157)
CDS                /gene="SH3BGR"
join(3473. .3565,12649. .12678,15331. .15426)
/note="Accession No. X93498"
/codon_start=1
/product="21-Glutamic Acid Rich protein 21-GARP"
/protein_id="CAB90445.1"
/db_xref="GI:7717370"
/db_xref="GOA:P55822"
/translation="GSEKAEKGTEAOKESGSDVGNLPEAQEKNEERGETATEETE
IAMEGAEAEAREEETASGERPGEDEDS"
3473. .3565
/gene="SH3BGR"
/number=4
repeat_region      /note="AluSx"
5542. .5829
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
repeat_region      complement(6887. .7067)
/note="LiMD3"
/rpt_family="LINE/L1"
/rpt_type=DISPERSED
repeat_region      complement(8482. .8762)
/note="AluJo"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
repeat_region      complement(8865. .8964)
/note="L2"
/rpt_family="LINE/L2"
/rpt_type=DISPERSED
repeat_region      9131. .9424
/note="AluSg"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
repeat_region      complement(9723. .9882)
/note="L2"
/rpt_family="LINE/L2"
/rpt_type=DISPERSED
repeat_region      complement(9989. .10077)
/note="L2"
/rpt_family="LINE/L2"
/rpt_type=DISPERSED
repeat_region      complement(10078. .10429)
/note="THE18"
/rpt_family="LTR/MaLR"
/rpt_type=DISPERSED
repeat_region      complement(10430. .10651)
/note="L2"
/rpt_family="LINE/L2"
/rpt_type=DISPERSED
repeat_region      10980. .11000
/note="(TTTA)n"
/rpt_family="Simple_repeat"
/rpt_type=TANDEM
repeat_region      complement(11002. .11282)
/note="AluSp"

```

```

repeat_region      /rpt_family="SINE/Alu"
complement(11315. .11556)
/note="AluJo"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
repeat_region      /note="MIR"
11702. .11753
/rpt_family="SINE/MIR"
/rpt_type=DISPERSED
exon               12649. .12678
/gene="SH3BGR"
/number=5
repeat_region      complement(13115. .13425)
/note="AluSx"
/rpt_family="SINE/Alu"
/rpt_type=DISPERSED
repeat_region      complement(13598. .13892)

Query Match      100.0%; Score 584; DB 9; Length 340000;
Best Local Similarity 100.0%; Pred. No. 7.3e-151;
Matches 584; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGACATTATCCAGAAGATTTCCTAGACGTCTATTACAACTGACCTGAGACCATTGA 60
DB 163805 GGGACATTATCCAGAAGATTTCCTAGACGTCTATTACAACTGACCTGAGACCATTGA 163864
QY 61 TGGGCATAGAAATGGTCCATCGCTTTTGTCTCAGGGGGCGTTTGTGATGAAACAGACT 120
DB 163865 TGGGCATAGAAATGGTCCATCGCTTTTGTCTCAGGGGGCGTTTGTGATGAAACAGACT 163924
QY 121 CAGACATGTTTCATCAATGTTGACTATCTGACTGAACCTGCTTCTGAAGAAAACAGAACAA 180
DB 163925 CAGACATGTTTCATCAATGTTGACTATCTGACTGAACCTGCTTCTGAAGAAAACAGAACAA 163984
QY 181 CCAGGTTTTTCACCTGGCTTCTTGAACTCAATGAGTTTCCCATCAGGAGCCATTTCAGCA 240
DB 163985 CCAGGTTTTTCACCTGGCTTCTTGAACTCAATGAGTTTCCCATCAGGAGCCATTTCAGCA 164044
QY 241 AGTGGTTTTGTCAGTAAATCTGAATATCCGTGGGACAGTACCACCATTCCTGCTCCGGCA 300
DB 164045 AGTGGTTTTGTCAGTAAATCTGAATATCCGTGGGACAGTACCACCATTCCTGCTCCGGCA 164104
QY 301 CCGGCTACGTGTTTTTCTGGGACGTGGCGAGTCAAGTGTACAATGTCTTCCAAGAGCGTCC 360
DB 164105 CCGGCTACGTGTTTTTCTGGGACGTGGCGAGTCAAGTGTACAATGTCTTCCAAGAGCGTCC 164164
QY 361 CATACATTAAACTGGAAGACGCTGTTTGTGGGGCTCTGCTCGAAAGGCTGAACATCAGAT 420
DB 164165 CATACATTAAACTGGAAGACGCTGTTTGTGGGGCTCTGCTCGAAAGGCTGAACATCAGAT 164224
QY 421 TGGAGGAGCTCCACTCCCGCCGACCTTTTTCAGGGGGCTTACGCTTCTCCGTTATGCC 480
DB 164225 TGGAGGAGCTCCACTCCCGCCGACCTTTTTCAGGGGGCTTACGCTTCTCCGTTATGCC 164284
QY 481 TCTTCAGGAGATCGTGGCTCGCCACTTCATCAAGCCCTCGGACTCTCTTGAGACTACTGGC 540
DB 164285 TCTTCAGGAGATCGTGGCTCGCCACTTCATCAAGCCCTCGGACTCTCTTGAGACTACTGGC 164344
QY 541 AGGCTCTAGAAATTCGGGGGGGGAAGATTGTCGCGCTGTCTGA 584
DB 164345 AGGCTCTAGAAATTCGGGGGGGGAAGATTGTCGCGCTGTCTGA 164388

```

```

RESULT 9
HSA6078          HSA6078          933 bp      DNA      linear      PRI 11-MAY-2000
LOCUS            Homo sapiens beta3gal-T6 gene.
DEFINITION       Homo sapiens beta3gal-T6 gene.
ACCESSION        AJ006078
VERSION          AJ006078.1 GI:7799922
KEYWORDS          beta-1,3-galactosyltransferase; beta3gal-T6 gene.
SOURCE            Homo sapiens (human)
ORGANISM          Homo sapiens

```


ORIGIN	assembly_fragment-clone_end.SP6-vector_side:right"																
Query Match	97.5%; Score 569.6; DB 9; Length 192219;																
Best Local Similarity	98.5%; Pred. No. 7.2e-147;																
Matches 575; Conservative	0; Mismatches 9; Indels 0; Gaps 0;																
Qy	1	GGGACATTAATCCAGAAAGGATTTCTTAGAGCTCTATTATCAATCTGACCCCTGAAGACCATGA	60														
Db	87020	GGGACATCAATCCAGAAGGATTTCTCTGGAGCTCTATTATCAATCTGACCCCTGAAGACCATGA	87079														
Qy	61	TGGGCGATAGAAATGGGTCCATCGCTTTTGTCCCTCAGCGCGGCTTTCTGATGAAGAACAGACT	120														
Db	87080	TGGGCGATAGAAATGGGTCCATCGCTTTTGTCCCTCAGCGCGGCTTTGTGATGAAGAACAGACT	87139														
Qy	121	CAGACATGTTTCATCAATGTTTGACTATCTGACTGAACTGCTTCTGAAGAAAAACAGAACAA	180														
Db	87140	CAGACATGTTTCATCAATGTTTGACTATCTGACTGAACTGCTTCTGAAGAAAAACAGAACAA	87199														
Qy	181	CCAGGTTTTTCACTGGCTTCTTTGAAATCAATGAGTTCCTCCATCAGGAGGCATTCAGCA	240														
Db	87200	CCAGGTTTTTCACTGGCTTCTTTGAAATCAATGAGTTCCTCCATCAGGAGGCATTCAGTA	87259														
Qy	241	AGTGGTTTGTTCAGTAAATCTGAATATCCGTGGGACAGGTACCCACCATTCTGCTCCGGCA	300														
Db	87260	AGTGGTTTGTTCAGTAAATCTGAATATCCGTGGGACAGGTACCCGCCATTCTGCTCCGGCA	87319														
Qy	301	CCGGCTTACGTGTTTCTGCGGACGCTGGCGAGTGCAGGTGCAGTGTACAACTCTCTCCAAGAGCGTCC	360														
Db	87320	CCGGCTTACGTGTTTCTGCGGACGCTGGCGAGTGCAGGTGTACAACTCTCTCCGAGAGCGTCC	87379														
Qy	361	CATACATTAAACTCGGAAGACGCTGTTTGTGGGCTCTGCCTCGAAAGGCTGAACATCAGAT	420														
Db	87380	CATACATTAAACTCGGAAGACGCTGTTTGTGGGCTCTGCCTCGAAAGGCTGAACATCAGAT	87439														
Qy	421	TGGAGAGCTCCACTCCGACGCGACTTTTTCAGGGGGCTTACGCTTCTCCGTATGCC	480														
Db	87440	TGGAGAGCTCCACTCCGACGCGACTTTTTCAGGGGGCTTACGCTTCTCCGTATGCC	87499														
Qy	481	TCCTTCAGGAGGATCGTGGGCTGCCACTTCATCAAGCCCTCGGACTCTCTTGGACTACTGGC	540														
Db	87500	GCCTTCAGGAGGATCGTGGGCTGCCACTTCATCAAGCCCTCGGACTCTCTTGGACTACTGGC	87559														
Qy	541	AGGCTCTAGAGAAATCCCGGGGGGAGATTTGTCGCCCTGTCTGA	584														
Db	87560	AGGCTCTAGAGAAATCCCGGGGGGAGATTTGTCCGCTATCTGA	87603														
RESULT 12																	
AB041416																	
LOCUS	AB041416 Homo sapiens betal,3-GalT 5 gene for UDP-Gal:GlcNAc																
DEFINITION	betal,3-galactosyltransferase 5, partial cds.																
ACCESSION	AB041416																
VERSION	AB041416.1 GI:7593026																
KEYWORDS																	
SOURCE	Homo sapiens (human)																
ORGANISM	Homo sapiens																
	Eukaryota; Metazoa; Ch																

```

source
1. .1576
/organism="Homo sapiens"
/mol_type="genomic DNA"
/isolate="#056"
/db_xref="taxon:9606"
/notes="human sequence used for primer design based on Acc#
AB020337"
<1. .27
/number=3
28. .674
/number=3
675. .1576
/genes="beta1,3-Galt 5"
675. .>1576
/genes="beta1,3-Galt 5"
/codon_start=1
/product="UDP-Gal:GlcNAc beta1,3-galactosyltransferase 5"
/protein_id="BAA94501.1"
/db_xref="GI:7593027"
/translation="WAPKRMVLYICLLVGLALCLYFSMYSNLNPKQSFVYKDGNF
LKLPTDCRQTPPELVLVTSKQLAERMAIRQWKGRTVKGKQLKTFLLGTSS
AETKEVDQESQRHGDIIQKDFLDYVNLTKTMGIEWHRCFQPAAFVMTKDSMF
INVDYLTKLKNRTTRPFTGLKNEPIROPFSKWFVSKSEYPMWRPFPFCSGVG
YVFGDVASQVYVNSKSPYIKLEDVFGVGLCLERLNIRLELHSHQPTFFPGGLRFSVC
LFRRIACHFIKPRTLDDYWQALE"
675. .>1576
/genes="beta1,3-Galt 5"
/number=4

ORIGIN
Query Match 94.7%; Score 553; DB 9; Length 1576;
Best Local Similarity 100.0%; Pred. No. 2.7e-142; Indels 0; Gaps 0;
Matches 553; Conservative 0; Mismatches 0;

QY 1 GGGACATTATCCAGAGGATTTCTTAGACGCTATTACAATCTGACCTGAAGACCATGA 60
DB 1024 GGGACATTATCCAGAGGATTTCTTAGACGCTATTACAATCTGACCTGAAGACCATGA 1083
QY 61 TGGGCATAGATGGTTCATCGCTTTTCTCAGCGCGGCTTTGTGATGAACAGACT 120
DB 1084 TGGGCATAGATGGTTCATCGCTTTTCTCAGCGCGGCTTTGTGATGAACAGACT 1143
QY 121 CAGCATGTTTCATCAATGTTGACTATCTGACTGAACTCAATGAGTTTCCCAAGAAAACAGAACAA 180
DB 1144 CAGCATGTTTCATCAATGTTGACTATCTGACTGAACTCAATGAGTTTCCCAAGAAAACAGAACAA 1203
QY 181 CCAGGTTTTTCACTGGCTTCTTGAAACTCAATGAGTTTCCCAAGAAAACAGAACAA 240
DB 1204 CCAGGTTTTTCACTGGCTTCTTGAAACTCAATGAGTTTCCCAAGAAAACAGAACAA 1263
QY 241 AGTGGTTTGTGAGTAAATCTGAATATCGTGGGACAGGTACCCACCATTTCTGCTCCGGCA 300
DB 1264 AGTGGTTTGTGAGTAAATCTGAATATCGTGGGACAGGTACCCACCATTTCTGCTCCGGCA 1323
QY 301 CCGGCTAGCTGTTTCTGGCGACGTGGCGAGTCAGGTGTACAATGTTCTCCCAAGAGCGTCC 360
DB 1324 CCGGCTAGCTGTTTCTGGCGACGTGGCGAGTCAGGTGTACAATGTTCTCCCAAGAGCGTCC 1383
QY 361 CATACATTAACATGGAAGACGTTGTTGTGGGGCTCTGCTCGAAAGGCTGAACATCAGAT 420
DB 1384 CATACATTAACATGGAAGACGTTGTTGTGGGGCTCTGCTCGAAAGGCTGAACATCAGAT 1443
QY 421 TGGAGAGCTCCACTCCAGCGGACCTTTTTCAGGGGGCTTACGCTTCTCCGTATGCC 480
DB 1444 TGGAGAGCTCCACTCCAGCGGACCTTTTTCAGGGGGCTTACGCTTCTCCGTATGCC 1503
QY 481 TCTTCAGAGGATCGTGGCCTGCCACTTTCATCAAGCCTCGAGCTCTCTTTGAGCTACTGGC 540
DB 1504 TCTTCAGAGGATCGTGGCCTGCCACTTTCATCAAGCCTCGAGCTCTCTTTGAGCTACTGGC 1563
QY 541 AGGCTCTAGAGAA 553
DB 1564 AGGCTCTAGAGAA 1576

```

```

RESULT 13
AB041415
LOCUS
DEFINITION
Pan paniscus beta1,3-Galt 5 gene for UDP-Gal:GlcNAc
beta1,3-galactosyltransferase 5, partial cds.
AB041415
ACCESSION
AB041415.1 GI:7593024
VERSION
KEYWORDS
SOURCE
Pan paniscus (pygmy chimpanzee)
ORGANISM
Pan paniscus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE
1 (bases 1 to 1570)
Liu, Y. and Saitou, N.
Silver Project
Published Only in DataBase (2000)
2 (bases 1 to 1570)
Liu, Y. and Saitou, N.
Direct Submission
Submitted (11-Apr-2000) Naryya Saitou, National Institute of
Genetics, Laboratory of Evolutionary Genetics; Ilii Yata, Mishima,
Shizuoka 411-8540, Japan (E-mail:nsaitou@genes.nig.ac.jp,
URL:http://sayer.lab.nig.ac.jp/~silver/, Tel:81-559-81-6790,
Fax:81-559-81-6789)
FEATURES
Location/Qualifiers
1. .1570
/organism="Pan paniscus"
/mol_type="genomic DNA"
/isolate="bonobo-05"
/db_xref="taxon:9597"
/notes="human sequence used for primer design based on Acc#
AB020337"
<1. .20
/number=3
21. .667
/number=3
668. .1570
/genes="beta1,3-Galt 5"
668. .>1570
/genes="beta1,3-Galt 5"
/codon_start=1
/product="UDP-Gal:GlcNAc beta1,3-galactosyltransferase 5"
/protein_id="BAA94500.1"
/db_xref="GI:7593025"
/translation="MAPPKRMVLYICLLVGLALCLYFSMYSNLNFKEQSFVYKDGNF
LKLPTDCRQTPPELVLVTSKQLAERMAIRQWKGRTVKGKQLKTFLLGTSS
AETKEVDQESQRHGDIIQKDFLDYVNLTKTMGIEWHRCFQPAAFVMTKDSMF
INVDYLTKLKNRTTRPFTGLKNEPIROPFSKWFVSKSEYPMWRPFPFCSGVG
YVFGDVASQVYVNSVSPYIKLEDVFGVGLCLERLNIRLELHSHQPTFFPGGLRFSVC
LFRRIACHFIKPRTLDDYWQALE"
668. .>1570
/genes="beta1,3-Galt 5"
/number=4

ORIGIN
Query Match 92.7%; Score 541.2; DB 9; Length 1570;
Best Local Similarity 98.6%; Pred. No. 5.1e-139;
Matches 546; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 GGGACATTATCCAGAGGATTTCTTAGACGCTATTACAATCTGACCTGAAGACCATGA 60
DB 1017 GGGACATTATCCAGAGGATTTCTTAGACGCTATTACAATCTGACCTGAAGACCATGA 1076
QY 61 TGGGCATAGATGGTTCATCGCTTTTGTCTCAGCGCGGCTTTGTGATGAACAGACT 120
DB 1077 TGGGCATAGATGGTTCATCGCTTTTGTCTCAGCGCGGCTTTGTGATGAACAGACT 1136
QY 121 CAGCATGTTTCATCAATGTTGACTATCTGACTGAACTCAATGAGTTTCCCAAGAAAACAGAACAA 180
DB 1137 CAGCATGTTTCATCAATGTTGACTATCTGACTGAACTCAATGAGTTTCCCAAGAAAACAGAACAA 1196

```

RESULT 15	AB041413	1565 bp	DNA	linear	PRI 13-APR-2000
LOCUS	AB041413				
DEFINITION	Homo sapiens beta1,3-GalT 5 gene for UDP-Gal:GlcNAc beta1,3-galactosyltransferase 5, partial cds.				
ACCESSION	AB041413				
VERSION	AB041413.1	GI:7593020			
KEYWORDS	.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 1565)				
AUTHORS	Liu,Y. and Saitou,N.				
TITLE	Silver Project				
JOURNAL	Published Only in DataBase (2000)				
REFERENCE	2 (bases 1 to 1565)				
AUTHORS	Liu,Y. and Saitou,N.				
TITLE	Direct Submission				

JOURNAL

Submitted (11-APR-2000) Naruya Saitou, National Institute of
Genetics, Laboratory of Evolutionary Genetics; 1111 Yata, Mishima,
Shizuoka 411-8540, Japan (E-mail:nsaitou@genes.nig.ac.jp,
URL: http://sayer.iab.nig.ac.jp/~silver/, Tel: 81-559-81-6790,
Fax: 81-559-81-6789)
Location/Qualifiers

FEATURES
source

1. .1565
/organism="Homo sapiens"
/mol_type="genomic DNA"
/isolate="human-NR"
/db_xref="taxon:9606"
/note="human sequence used for primer design based on Acc#
AB020337"
<1. .27
/number=3
28. .674
/number=3
675. .1565
/gene="beta1,3-GalT 5"
675. .>1565
/gene="beta1,3-GalT 5"
/codon_start=1
/product="UDP-Gal:GlcNAc beta1,3-galactosyltransferase 5"
/protein_id="BAA94498.1"
/db_xref="GI:7593021"
/translation="MAPPKRLMVLCLVLGALCLFYSWYSLNPFKEQSFVYKKGDNF
LKLPTDQRQTPPELVLLVTSKQLAERMAIRQTWGERXVKGKQLKTFPLLGTTSS
AAETKEVDQESQRGDIIQKDFLDVYINLTAKTMGIEWHRFCQAAPVNTKTDNDMP
INVLYLLELLKKNRTTRFFFLKLINEFFPIRQPFKSWFVSKSEYPWDYRPPPCSGTG
YVFSGDVASQVYVSKVPIYKLEDFVFLGCLERLNIRLELHSQPTFFPFGGLRFSVC
LFRIVACHFIKPRTLMTTGR"
675. .>1565
/gene="beta1,3-GalT 5"
/number=4

ORIGIN

Query Match 90.9%; Score 531; DB 9; Length 1565;
Best Local Similarity 99.8%; Pred No. 3.4e-136;
Matches 542; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 GGGACATTATCCAGAAGGATTTCCTAGACGCTATTACAATCTGACCCCTGAAGACCATGA 60
Db 124 GGGACATTATCCAGAAGGATTTCCTAGACGCTATTACAATCTGACCCCTGAAGACCATGA 1083
QY 61 TGGGCATAGATGGTCCATCGCTTTTGTCTCAGCGCGCGCTTTGTGATGAACACAGACT 120
Db 184 TGGGCATAGATGGTCCATCGCTTTTGTCTCAGCGCGCGCTTTGTGATGAACACAGACT 1143
QY 121 CAGACATGTTTCATCAATGTTGACTATCTGACTGAACTGCTTTCTGAAGAAAAACAGAACAA 180
Db 144 CAGACATGTTTCATCAATGTTGACTATCTGACTGAACTGCTTTCTGAAGAAAAACAGAACAA 1203
QY 181 CCAGGTTTTTCATCGGCTTCTGAACTCAATGAGTTTCCATCAGGCGCCATTCAGCA 240
Db 104 CCAGGTTTTTCATCGGCTTCTGAACTCAATGAGTTTCCATCAGGCGCCATTCAGCA 1263
QY 241 AGTGGTTTGTGAGTAAATCTGAATATCCGTGGGACAGGTACCCCACTTCTGCTCCGGCA 300
Db 164 AGTGGTTTGTGAGTAAATCTGAATATCCGTGGGACAGGTACCCCACTTCTGCTCCGGCA 1323
QY 301 CCGGCTAGCTGTTTTCTGGCGACGTGGCGAGTCAAGTGTACAATGTCTCCAAGAGCGTCC 360
Db 124 CCGGCTAGCTGTTTTCTGGCGACGTGGCGAGTCAAGTGTACAATGTCTCCAAGAGCGTCC 1383
QY 361 CATACATTAAATCGAAGACGTGTTTGTGGGGCTCTGCTCGAAAGGCTGAACATCAGAT 420
Db 184 CATACATTAAATCGAAGACGTGTTTGTGGGGCTCTGCTCGAAAGGCTGAACATCAGAT 1443
QY 421 TGGAGGAGCTCCACTCCAGCGACCTTTTTTCCAGGGGGCTTAGCGTTCTCCGTATGCC 480
Db 144 TGGAGGAGCTCCACTCCAGCGACCTTTTTTCCAGGGGGCTTAGCGTTCTCCGTATGCC 1503
QY 481 TCTTCAGGAGGATCGTGGCTCGCCACTTTCATCAAGCCTCGGACTCTCTTGGACTACTGGC 540

Db 1504 TCTTCAGGAGGATCGTGGCTCGCCACTTTCATCAAGCCTCGGACTCTC-TGGACTACTGGC 1562
QY 541 AGG 543
Db 1563 AGG 1565

Search completed: September 22, 2005, 14:50:07
Job time : 2673.69 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2005, 05:52:47 ; Search time 348.351 Seconds
(without alignments)
9924.272 Million cell updates/sec

Title: US-10-777-828-8_COPY_428_1011

Perfect score: 584

Sequence: 1 gggacattatccagaagat.....aagattgtccgcgtgtctga 584

Scoring table: IDENTITY_NUC

Gapop 10_0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001as:*
5: Geneseq2001bs:*
6: Geneseq2002as:*
7: Geneseq2002bs:*
8: Geneseq2003as:*
9: Geneseq2003bs:*
10: Geneseq2003cs:*
11: Geneseq2003ds:*
12: Geneseq2004as:*
13: Geneseq2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	584	100.0	933	12 ADK68456	Adk68456 Human bet
2	584	100.0	1011	3 AA27959	Aa27959 Human bet
3	584	100.0	2629	10 ABZ56918	Abz56918 B3GALT nu
4	584	100.0	2762	10 ABZ56917	Abz56917 B3GALT nu
5	584	100.0	2775	3 AA93875	Aa93875 Human bet
6	584	100.0	3409	10 ABZ56919	Abz56919 B3GALT nu
7	584	100.0	10562	3 AA93876	Aa93876 Human bet
8	583	99.8	933	10 ABZ56916	Abz56916 B3GALT nu
9	114.4	19.6	1266	2 AA35710	Aa35710 cDNA enco
10	114.4	19.6	1268	10 ABZ56911	Abz56911 B3GALT nu
11	114.4	19.6	1269	8 AD48744	Ad48744 Murine be
12	114.4	19.6	1269	10 ABZ56913	Abz56913 B3GALT nu
13	114.4	19.6	1473	12 ADM66890	Adm66890 Murine ad
14	114.4	19.6	2420	2 AA35711	Aa35711 cDNA enco
15	114.4	19.6	3054	12 ADM66891	Adm66891 Human adi
16	114.4	19.6	3212	10 ABZ56914	Abz56914 B3GALT nu
17	114.4	19.6	3215	10 ABZ56912	Abz56912 B3GALT nu
18	114.4	19.6	3215	10 ABZ56915	Abz56915 B3GALT nu
19	110.6	18.9	1037	10 ABZ56909	Abz56909 B3GALT nu
20	110.6	18.9	1739	2 AAQ67067	Aaq67067 Beta-1,3-

21	109	18.7	2168	10 ABZ56910	Abz56910 B3GALT nu
22	106.2	18.2	1092	12 ADO00390	Ado00390 Novel hum
23	106.2	18.2	1092	12 ADN98821	Adn98821 Novel hum
24	106.2	18.2	1358	10 ADF76806	Adf76806 Novel hum
25	106.2	18.2	1773	2 AA87193	Aa87193 Human Den
26	106.2	18.2	1897	6 ABK51201	Abk51201 Human cDN
27	106.2	18.2	1897	12 ADO19801	Ado19801 Human PRO
28	106.2	18.2	2024	9 ADA20069	Ada20069 Novel hum
29	106.2	18.2	2095	3 AAZ65022	Aaz65022 Membrane-
30	106.2	18.2	2095	4 AAS46005	Aas46005 Human DNA
31	106.2	18.2	2095	4 AAF92075	Aaf92075 Human PRO
32	106.2	18.2	2095	5 AAF44168	Aaf44168 Human PRO
33	106.2	18.2	2095	6 ABS74395	Abs74395 Human cDN
34	106.2	18.2	2095	8 ACA89455	Aca89455 cDNA enco
35	106.2	18.2	2095	8 ACA73465	Aca73465 Human sec
36	106.2	18.2	2095	8 ACA05780	Aca05780 Human sec
37	106.2	18.2	2095	8 ACA66614	Aca66614 cDNA enco
38	106.2	18.2	2095	8 ACA64304	ACA64304 Novel hum
39	106.2	18.2	2095	8 ACA91181	Ac91181 Novel hum
40	106.2	18.2	2095	8 ACD81558	Ac81558 Human cDN
41	106.2	18.2	2095	8 ACF20189	Acf20189 Human sec
42	106.2	18.2	2095	8 ACF19575	Acf19575 Human sec
43	106.2	18.2	2095	8 ACD21863	Ac21863 Human sec
44	106.2	18.2	2095	8 ACF13028	Acf13028 Human sec
45	106.2	18.2	2095	8 ACD25131	Ac25131 Human sec

ALIGNMENTS

RESULT 1

ID ADK68456 standard; cDNA; 933 BP.

AC ADK68456;

DT 06-MAY-2004 (first entry)

XX Human beta-1,3-galactosyl transferase cDNA SeqID 1.

XX human; gene; ss; saccharide binding protein; maltose binding protein;
KW beta-1,3-galactosyl transferase; inflammation; infectious disease;
KW cancer metastasis suppression; dairy product; antiinflammatory;
KW antimicrobial; cytostatic.

OS Homo sapiens.

XX Key Location/Qualifiers

FH CDS 1..933

FT /*tag= a

FT /product= "Beta-1,3-galactosyl transferase protein"

XX JP2004016117-A.

XX 22-JAN-2004.

XX 17-JUN-2002; 2002JP-00176132.

XX 17-JUN-2002; 2002JP-00176132.

XX (TOYM) TOYOBO KK.

XX WPI; 2004-113878/12.

XX P-PSDB; ADK68457.

XX New fusion protein useful in diagnosis of diseases and in manufacture of pharmaceutical products, comprises a saccharide binding protein and beta 1, 3-galactosyl transferase.

XX Example 1; SEQ ID NO 1; 23pp; Japanese.

XX This invention relates to a novel recombinant fusion protein that

CC comprises a saccharide binding protein, in particular a maltose binding

CC protein, and a beta-1,3-galactosyl transferase and an appropriate
 CC manufacturing method. Specifically, it refers to fusion protein that can
 CC transfer galactose to an N-acetyl glucosamine residue or N-acetyl
 CC glucosamine monosaccharide. The present invention describes a method to
 CC produce this enzyme, cheaply and efficiently, for the diagnosis and
 CC treatment of inflammation, infectious diseases or for cancer metastasis
 CC suppression. Furthermore, it can also be useful for improving dairy
 CC products. Accordingly, compositions exhibit various activities including
 CC anti-inflammatory, antimicrobial and cytostatic. This polynucleotide
 CC sequence is the human beta-1,3-galactosyl transferase cDNA of the
 CC invention.

XX SQ Sequence 933 BP; 227 A; 232 C; 241 G; 233 T; 0 U; 0 Other;

Query Match 100.0%; Score 584; DB 12; Length 933;
 Best Local Similarity 100.0%; Pred. No. 5.1e-181;
 Matches 584; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGACATTATCCAGAGGATTTCTAGACGCTATTACAACTGACCTGGAAGACCATGA 60
 DB 356 GGGACATTATCCAGAGGATTTCTAGACGCTATTACAACTGACCTGGAAGACCATGA 409

QY 61 TGGGCATAGAAATGGGTCCATCGCTTTTCTCAGCGCGCGTTTGTGATGAAACAGACT 120
 DB 410 TGGGCATAGAAATGGGTCCATCGCTTTTCTCAGCGCGCGTTTGTGATGAAACAGACT 469

QY 121 CAGACATGTTCAATGATGACTATCTGACTGAACTGCTTCTGAGAAACAGAACAA 180
 DB 474 CAGACATGTTCAATGATGACTATCTGACTGAACTGCTTCTGAGAAACAGAACAA 529

QY 181 CCAGGTTTTTCACTGGCTTCTTGAACACTCAATGAGTTTCCCATCAGGCGAGCCATTTCAGCA 240
 DB 530 CCAGGTTTTTCACTGGCTTCTTGAACACTCAATGAGTTTCCCATCAGGCGAGCCATTTCAGCA 589

QY 241 AGTGTTTGTTCAGTAAATCTGAATATCCGTGGGACAGGTACCCACATTCGTCTCCGCA 300
 DB 590 AGTGTTTGTTCAGTAAATCTGAATATCCGTGGGACAGGTACCCACATTCGTCTCCGCA 649

QY 601 CCGGTACGTGTTTCTCGGCGAGTGGGAGTGTGACATGCTCTCCAGAGCGTCC 360
 DB 650 CCGGTACGTGTTTCTCGGCGAGTGGGAGTGTGACATGCTCTCCAGAGCGTCC 709

QY 661 CATACATTAAACTGGAAGACGCTGTTTGTGGGCTCTGCTCTGAAAGGCTGAACATCAGAT 420
 DB 710 CATACATTAAACTGGAAGACGCTGTTTGTGGGCTCTGCTCTGAAAGGCTGAACATCAGAT 769

QY 421 TGGAGGAGCTCCACTCCCGAGCGACCTTTTTCAGGGGGCTTACGCTTCTCCGATGCC 480
 DB 170 TGGAGGAGCTCCACTCCCGAGCGACCTTTTTCAGGGGGCTTACGCTTCTCCGATGCC 829

QY 481 TCTTCAGGAGGATCGTGGCTGCGACCTTCATCAGCCCTCGGACTCTTGGACTACTGGC 540
 DB 830 TCTTCAGGAGGATCGTGGCTGCGACCTTCATCAGCCCTCGGACTCTTGGACTACTGGC 889

QY 541 AGGCTCTAGAGAAATCCCGGGGGGAAGATTGTGCGCCCTGTCTGA 584
 DB 890 AGGCTCTAGAGAAATCCCGGGGGGAAGATTGTGCGCCCTGTCTGA 933

RESULT 2
 ID AAA27959
 XX AAA27959 standard; DNA; 1011 BP.
 AC AAA27959
 XX
 DT 15-AUG-2000 (first entry)
 XX
 DE Human beta3Gal-T5 gene sequence.
 XX
 KW UDP-D-galactose; beta-N-acetylglucosamine beta1,3-galactosyltransferase;
 KW Beta3Gal-T5; ss; human; chromosome 21q22.3; galactosylation;
 KW beta1,3-galactosyl glycosylated saccharide production; glycopeptide;
 KW glycoprotein.

XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH CDS 79..1011
 FT /*tag= a
 FT /product= "Beta3Gal-T5"
 FT complement(79..98)
 FT primer_bind /*tag= b
 FT primer_bind complement(150..170)
 FT primer_bind /*tag= c
 FT primer_bind 991..1011
 FT primer_bind /*tag= d
 XX
 XX WO200029558-A1.
 XX 25-MAY-2000.
 XX 11-NOV-1999; 99WO-US026807.
 XX 13-NOV-1998; 98DK-00001483.
 XX (CLAU/) CLAUSEN H.
 XX Clausen H, Amado M;
 XX WPI; 2000-399728/34.
 XX P-PSDB; AAY94641.
 XX Novel nucleic acid sequence encoding human UDP-galactose:beta-N-
 XX acetylglucosamine beta1,3-galactosyltransferase useful for obtaining beta
 XX 1,3-galactosyl glycosylated saccharides and glycopeptides or
 XX glycoproteins.
 XX
 XX Claim 7; Fig 1; 74pp; English.
 XX
 XX The present invention relates to a nucleic acid sequence encoding UDP-D-
 XX galactose:beta-N-acetylglucosamine beta1,3-galactosyltransferase
 XX (beta3Gal-T5). Beta3 transferases add galactose to the hydroxy group at
 XX carbon 3 of 2-acetamido-2-deoxy-D-glucose (GlcNAc). The present sequence
 XX represents the human beta3Gal-T5 gene sequence. The beta3Gal-T5 gene is
 XX located on human chromosome 21q22.3. Beta3Gal-T5 is a type II
 XX transmembrane glycoprotein. The invention also relates to the beta3Gal-T5
 XX protein sequence, a nucleic acid vector comprising the beta3Gal-T5
 XX nucleotide sequence, a host cell comprising the vector, and a method for
 XX the production of the beta3Gal-T5 protein from the host cells. The
 XX methods of the invention can be used for recombinant production of
 XX beta3Gal-T5 for use as a catalyst and for recombinant production of
 XX peptides or proteins with appropriate galactosylation. The beta3Gal-T5
 XX protein can be used to obtain beta1,3-galactosyl glycosylated
 XX saccharides, glycopeptides or glycoproteins

XX SQ Sequence 1011 BP; 247 A; 256 C; 251 G; 257 T; 0 U; 0 Other;
 Query Match 100.0%; Score 584; DB 3; Length 1011;
 Best Local Similarity 100.0%; Pred. No. 5.3e-181;
 Matches 584; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGACATTATCCAGAGGATTTCTAGACGCTTATTACAACTGACCTGGAAGACCATGA 60
 DB 428 GGGACATTATCCAGAGGATTTCTAGACGCTTATTACAACTGACCTGGAAGACCATGA 487

QY 61 TGGGCATAGAAATGGGTCCATCGCTTTTGTCTCAGGGGGCTTGTGATGAAACAGACT 120
 DB 488 TGGGCATAGAAATGGGTCCATCGCTTTTGTCTCAGGGGGCTTGTGATGAAACAGACT 547

QY 121 CAGACATGTTCAATGATGACTATCTGACTGAACTGCTTCTGAGAAACAGAACAA 180
 DB 548 CAGACATGTTCAATGATGACTATCTGACTGAACTGCTTCTGAGAAACAGAACAA 607

QY 181 CCAGGTTTTTCACTGGCTTCTTGAACACTCAATGAGTTTCCCATCAGGCGAGCCATTTCAGCA 240
 DB 608 CCAGGTTTTTCACTGGCTTCTTGAACACTCAATGAGTTTCCCATCAGGCGAGCCATTTCAGCA 667

QY 241 AGTGGTTTCTCAGTAAATCTGAATATCCCTGGGACAGGTACCCACCAATTCCTCTCCGGCA 300
DB 668 AGTGGTTTCTCAGTAAATCTGAATATCCCTGGGACAGGTACCCACCAATTCCTCTCCGGCA 727
QY 301 CCGGCTACGTGTTTCTGGGACGCTGGGAGTCAGGTGTACAAATCTCTCCAGAGCGTCC 360
DB 728 CCGGCTACGTGTTTCTGGGACGCTGGGAGTCAGGTGTACAAATCTCTCCAGAGCGTCC 787
QY 361 CATACATTAATCTGGAAGACGTGTTTGTGGGCTCTGCCCTCGAAGGCTGAACATCAGAT 420
DB 788 CATACATTAATCTGGAAGACGTGTTTGTGGGCTCTGCCCTCGAAGGCTGAACATCAGAT 847
QY 421 TGGAGAGCTCCACTCCAGCCGACCTTTTTCAGGGGGCTTACGCTTCTCCGTATGCC 480
DB 848 TGGAGAGCTCCACTCCAGCCGACCTTTTTCAGGGGGCTTACGCTTCTCCGTATGCC 907
QY 481 TCTTCAGGAGGATCGTGGCTGCCACTTTCATCAAGCCTCGGACTCTCTTGGACTACTGCC 540
DB 908 TCTTCAGGAGGATCGTGGCTGCCACTTTCATCAAGCCTCGGACTCTCTTGGACTACTGCC 967
QY 541 AGGCTCTAGAGAAATCCCGGGGGGAAGATTGTCGCTGTCTGA 584
DB 968 AGGCTCTAGAGAAATCCCGGGGGGAAGATTGTCGCTGTCTGA 1011

RESULT 3

ID ABZ56918 standard; DNA; 2629 BP.
XX AC ABZ56918;
XX AC ABZ56918;
DT 04-APR-2003 (first entry)
XX B3GALT nucleic acid sequence # SEQ ID 10.
XX B3GALT; beta-1,3 galactosyltransferase; p53; cancer; breast; colon;
KW kidney; lung; ovary; gene; ds.
XX Homo sapiens.
FN WO200299044-A2.
XX PD 12-DEC-2002.
XX PF 02-JUN-2002; 2002WO-US017356.
XX PR 05-JUN-2001; 2001US-0296076P.
PR 10-OCT-2001; 2001US-0328605P.
PR 15-FEB-2002; 2002US-0357253P.
XX PA (EXEL-) EXELIXIS INC.
XX PI Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
XX WPI; 2003-156849/15.

DR Identifying p53 pathway modulating agents with B3GALT genes, useful for
PT the diagnosis and treatment of disorders associated with defects in the
PT p53 pathway, such as cancer of the breast, colon, kidneys, lung and
PT ovary.
XX Disclosure; Page 53-55; 82pp; English.
XX The invention relates to identifying a candidate p53 pathway modulating
CC agent in humans that is referred to in the specification as BGALT (beta-
CC 1,3 galactosyltransferase). Also disclosed is a method for diagnosing a
CC disease in a patient, by contacting a sample with a probe for B3GALT
CC expression, and comparing the results with a control, and determining
CC whether the results indicate a likelihood of disease. Methods and
CC compositions of the invention are useful for the diagnosis and treatment
CC of disorders associated with defects in the p53 pathway, such as cancer
CC of the breast, colon, kidneys, lung and ovary. The current sequence

CC represents a B3GALT nucleic acid sequence referred to in the disclosure
CC of the invention
XX SQ Sequence 2629 BP; 632 A; 665 C; 644 G; 688 T; 0 U; 0 Other;
Query Match 100.0%; Score 584; DB 10; Length 2629;
Best Local Similarity 100.0%; Pred. No. 9e-181;
Matches 584; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGACATTTATCCAGAAAGGATTTCTTAGACGTCTATTACAACTGACCCCTGAAGACCATCA 60
DB 650 GGGACATTTATCCAGAAAGGATTTCTTAGACGTCTATTACAACTGACCCCTGAAGACCATCA 709
QY 61 TGGGCATAGAATGGGTCCATCGCTTTTCTCAGCGCGCGTTTGTGATGAAACAGACT 120
DB 710 TGGGCATAGAATGGGTCCATCGCTTTTCTCAGCGCGCGTTTGTGATGAAACAGACT 769
QY 121 CAGACATGTTTATCAATGTTGACTATCTGACTGAATCTGTTCTGAAGAAAAACAGAA 180
DB 770 CAGACATGTTTATCAATGTTGACTATCTGACTGAATCTGTTCTGAAGAAAAACAGAA 829
QY 181 CCAGGTTTTTCACTGGCTTCTTGAACCTCAATGAGTTTCCCATCAGGCGACCATTCAGCA 240
DB 830 CCAGGTTTTTCACTGGCTTCTTGAACCTCAATGAGTTTCCCATCAGGCGACCATTCAGCA 889
QY 241 AGTGGTTTGTTCAGTAAATCTGAATATCCGTGGGACAGGTACCCACCATTTCTGCTCCGGCA 300
DB 890 AGTGGTTTGTTCAGTAAATCTGAATATCCGTGGGACAGGTACCCACCATTTCTGCTCCGGCA 949
QY 301 CCGGCTACGTGTTTCTGGCGACGCTGGGAGTCAATGTTCTCCAAAGAGCGTCC 360
DB 950 CCGGCTACGTGTTTCTGGCGACGCTGGGAGTCAATGTTCTCCAAAGAGCGTCC 1009
QY 361 CATACATTAATCTGGAAGACGTGTTTGTGGGGCTCTGCTCGAAGGCTGAACATCAGAT 420
DB 1010 CATACATTAATCTGGAAGACGTGTTTGTGGGGCTCTGCTCGAAGGCTGAACATCAGAT 1069
QY 421 TGGAGGAGCTCCACTCCAGCCGACCTTTTTCAGGGGGGCTTACGCTTCTCCGTATGCC 480
DB 1070 TGGAGGAGCTCCACTCCAGCCGACCTTTTTCAGGGGGGCTTACGCTTCTCCGTATGCC 1129
QY 481 TCTTCAGGAGGATCGTGGCCTGCCACTTTCATCAAGCCTCGGACTCTCTTGGACTACTGCC 540
DB 1130 TCTTCAGGAGGATCGTGGCCTGCCACTTTCATCAAGCCTCGGACTCTCTTGGACTACTGCC 1189
QY 541 AGGCTCTAGAGAAATCCCGGGGGGAAGATTGTCGCTGTCTGA 584
DB 1190 AGGCTCTAGAGAAATCCCGGGGGGAAGATTGTCGCTGTCTGA 1233

RESULT 4

ABZ56917
ID ABZ56917 standard; DNA; 2762 BP.
XX AC ABZ56917;
XX AC ABZ56917;
DT 04-APR-2003 (first entry)
XX DE B3GALT nucleic acid sequence # SEQ ID 9.
XX B3GALT; beta-1,3 galactosyltransferase; p53; cancer; breast; colon;
KW kidney; lung; ovary; gene; ds.
XX Homo sapiens.
XX WO200299044-A2.
XX PD 12-DEC-2002.
XX PF 02-JUN-2002; 2002WO-US017356.
XX PR 05-JUN-2001; 2001US-0296076P.
PR 10-OCT-2001; 2001US-0328605P.

PR 15-FEB-2002; 2002US-0357253P.
XX (EXEL-) EXELIXIS INC.
XX Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
XX WPI; 2003-156849/15.
XX
XX Identifying p53 pathway modulating agents with B3GALT genes, useful for
XX the diagnosis and treatment of disorders associated with defects in the
XX p53 pathway, such as cancer of the breast, colon, kidneys, lung and
XX ovary.
XX
XX Example 5; Page 52-53; 82pp; English.
XX
XX The invention relates to identifying a candidate p53 pathway modulating
XX agent in humans that is referred to in the specification as BGALT (beta-
XX 1,3 galactosyltransferase). Also disclosed is a method for diagnosing a
XX disease in a patient, by contacting a sample with a probe for B3GALT
XX expression, and comparing the results with a control, and determining
XX whether the results indicate a likelihood of disease. Methods and
XX compositions of the invention are useful for the diagnosis and treatment
XX of disorders associated with defects in the p53 pathway, such as cancer
XX of the breast, colon, kidneys, lung and ovary. The current sequence
XX represents a B3GALT nucleic acid sequence referred to in an example from
XX the invention
XX
XX Sequence 2762 BP; 651 A; 703 C; 679 G; 729 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 584; DB 10; Length 2762;
XX Best Local Similarity 100.0%; Pred. No. 9.2e-181;
XX Matches 584; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 GGGACATTATCCAGAGGATTTCCTAGACGCTCTATTACAATCTGACCTGAAGACCATGA 60
DB 763 GGGACATTATCCAGAGGATTTCCTAGACGCTCTATTACAATCTGACCTGAAGACCATGA 842
XX
QY 61 TGGGCATAGATGGTCCATCGCTTTCTCTCAGGGCGCTTTGTGATGAACAGACT 120
DB 843 TGGGCATAGATGGTCCATCGCTTTCTCTCAGGGCGCTTTGTGATGAACAGACT 902
XX
QY 121 CAGACATGTTTCATCAATCTTGACTATCTGACTGAACTCTTCTGAAGAAAACAGAACAA 180
DB 903 CAGACATGTTTCATCAATCTTGACTATCTGACTGAACTCTTCTGAAGAAAACAGAACAA 962
XX
QY 181 CCAGGTTTTTCATCGGCTTTCTGAAACTCAATAGTCTTCCATCAGGAGCCATTCAGCA 240
DB 963 CCAGGTTTTTCATCGGCTTTCTGAAACTCAATAGTCTTCCATCAGGAGCCATTCAGCA 1022
XX
QY 241 AGTGGTTTTGTGAGTAAATCTGAATATCCGTGGGACAGGTACCCACCTTCGCTCCGGCA 300
DB 1023 AGTGGTTTTGTGAGTAAATCTGAATATCCGTGGGACAGGTACCCACCTTCGCTCCGGCA 1082
XX
QY 301 CCGGCTAGCTGTTTTCTGCGGACGCTGGCGAGTCAAGTGTACAATGTCTCCAGAGCGTCC 360
DB 1083 CCGGCTAGCTGTTTTCTGCGGACGCTGGCGAGTCAAGTGTACAATGTCTCCAGAGCGTCC 1142
XX
QY 361 CATACATTAATCGAAGACGCTGTTGTGGGGCTCTGCTCTGAAAGGCTGAACATCAGAT 420
DB 1143 CATACATTAATCGAAGACGCTGTTGTGGGGCTCTGCTCTGAAAGGCTGAACATCAGAT 1202
XX
QY 421 TGGAGGAGCTCCACTCCAGCGGACCTTTTTCAGGGGGCTTACGCTTCTCCGATGCC 480
DB 1203 TGGAGGAGCTCCACTCCAGCGGACCTTTTTCAGGGGGCTTACGCTTCTCCGATGCC 1262
XX
QY 481 TCTTCAGGAGGATCGTGCGCTGCCACTTTCATCAAGCCTCGGACTCTCTTGAGCTACTGGC 540
DB 1263 TCTTCAGGAGGATCGTGCGCTGCCACTTTCATCAAGCCTCGGACTCTCTTGAGCTACTGGC 1322
XX
QY 541 AGGCTCTAGAGAAATCCCGGGGGGAGATGTGTCGCTGTCTGA 584
DB 1323 AGGCTCTAGAGAAATCCCGGGGGGAGATGTGTCGCTGTCTGA 1366

RESULT 5
AAA93875
ID AAA93875 standard; DNA; 2775 BP.
XX
XX AAA93875;
XX
XX 15-JAN-2001 (first entry)
XX
XX Human beta-1,3 galactose transferase encoding DNA.
XX
XX Beta-1,3 galactose transferase; treatment; diagnosis; cancer; human;
XX digestive system; ds.
XX
XX Homo sapiens.
XX
XX WO2000050608-A1.
XX
XX 31-AUG-2000.
XX
XX 24-FEB-2000; 2000WO-JP001070.
XX
XX 25-FEB-1999; 99JP-00047571.
XX
XX (KYOW) KYOWA HAKKO KOGYO KK.
XX
XX Narimatsu H, Ieshiki S, Togayachi A, Sasaki K;
XX
XX WPI; 2000-549409/50.
XX
XX P-PSDB; AAB93875.
XX
XX Beta-1,3 galactose transferase and DNA encoding it, useful for synthesis
XX of type 1 sialyl Lewis, a carbohydrate for treatment of digestive system
XX cancer.
XX
XX Claim 5; Page 99-102; 123pp; Japanese.
XX
XX This invention relates to a polypeptide (I) with beta-1,3 galactose
XX transferase activity, or variants of (I) comprising amino acid additions,
XX deletions and/or substitutions. Included in the invention is DNA encoding
XX all or part of (I); expression vectors containing the DNA, host cells
XX transformed by the vectors; a method for the preparation of the
XX polypeptide by culture of the transformants or by expression of the
XX of a transgenic mammal, and antibodies recognising (I). The Beta-1,3
XX galactose transferase protein transfers galactose by beta-1,3 bonding to
XX N-acetylglucosamine present in a non-cyclic carbohydrate chain (such as
XX GlcNAc) to give Galbeta1-4Glc to give Galbeta1-3GlcNAc. The protein and DNA
XX encoding it are useful for the treatment and diagnosis of cancer of the
XX digestive system. The present sequence represents Beta-1,3 galactose
XX transferase encoding DNA
XX
XX Sequence 2775 BP; 681 A; 698 C; 669 G; 727 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 584; DB 3; Length 2775;
XX Best Local Similarity 100.0%; Pred. No. 9.2e-181;
XX Matches 584; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 GGGACATTATCCAGAGGATTTCCTAGACGCTCTATTACAATCTGACCTGAAGACCATGA 60
DB 751 GGGACATTATCCAGAGGATTTCCTAGACGCTCTATTACAATCTGACCTGAAGACCATGA 810
XX
QY 61 TGGGCATAGATGGTCCATCGCTTTTGTCTCAGGGCGCTTTGTGATGAACAGACT 120
DB 811 TGGGCATAGATGGTCCATCGCTTTTGTCTCAGGGCGCTTTGTGATGAACAGACT 870
XX
QY 121 CAGACATGTTTCATCAATAGTGTGACTATCTGACTGAACCTGCTTCTGAAGAAAACAGAACAA 180
DB 871 CAGACATGTTTCATCAATAGTGTGACTATCTGACTGAACCTGCTTCTGAAGAAAACAGAACAA 930
XX
QY 181 CCAGGTTTTTCATCGGCTTCTTGAACTCAATAGTGTTCATCAGGAGCCATTCAGCA 240
DB 931 CCAGGTTTTTCATCGGCTTCTTGAACTCAATAGTGTTCATCAGGAGCCATTCAGCA 990

QY 241 AGTGGTTGTCAGTAAATCTGAATATCCGTGGGACAGGTACCCACATTTCTGCTCGGCA 300
Db |||||
991 AGTGGTTGTCAGTAAATCTGAATATCCGTGGGACAGGTACCCACATTTCTGCTCGGCA 1050
QY 301 CCGGCTACGTGTTTCTGCGGACGTGGCGAGTCAGGTGTACAAATCTCTCCAGAGCGTCC 360
Db |||||
1051 CCGGCTACGTGTTTCTGCGGACGTGGCGAGTCAGGTGTACAAATCTCTCCAGAGCGTCC 1110
QY 361 CATACATTAATCTGGAAGACGTGTTTGTGGGGCTCTGCTCGAAAGGCTGAACATCAGAT 420
Db |||||
1111 CATACATTAATCTGGAAGACGTGTTTGTGGGGCTCTGCTCGAAAGGCTGAACATCAGAT 1170
QY 421 TGGAGAGCTCCACCTCCAGCCGACCTTTTTCAGGGGGCTTACGCTTCTCCGTATGCC 480
Db |||||
1171 TGGAGAGCTCCACCTCCAGCCGACCTTTTTCAGGGGGCTTACGCTTCTCCGTATGCC 1230
QY 481 TCCTCAGGAGGATCGTGGGCTGCCACTTCATCAAGCCTCGGACTCTCTTGGACTACTGGC 540
Db |||||
1231 TCCTCAGGAGGATCGTGGGCTGCCACTTCATCAAGCCTCGGACTCTCTTGGACTACTGGC 1290
QY 541 AGGCTCTAGAGAAATCCCGGGGGGAAGATTGTCCGCTGTCTGA 584
Db |||||
1291 AGGCTCTAGAGAAATCCCGGGGGGAAGATTGTCCGCTGTCTGA 1334

RESULT 6

ABZ56919
ID ABZ56919 standard; DNA; 3409 BP.

XX AC ABZ56919;

XX DT 04-APR-2003 (first entry)

XX DE B3GALT nucleic acid sequence # SEQ ID 11.

XX KW B3GALT; beta-1,3 galactosyltransferase; p53; cancer; breast; colon;
XX LW kidney; lung; ovary; gene; ds.

XX OS Homo sapiens.

XX PN WO200299044-A2.

XX PD 12-DEC-2002.

XX PF 02-JUN-2002; 2002WO-US017356.

XX PR 05-JUN-2001; 2001US-0296076P.

XX PR 10-OCT-2001; 2001US-0328605P.

XX PR 15-FEB-2002; 2002US-0357253P.

XX PA (EXEL-) EXELIXIS INC.

XX PI Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;

XX DR WPI; 2003-156849/15.

XX PT Identifying p53 pathway modulating agents with B3GALT genes, useful for
XX PT the diagnosis and treatment of disorders associated with defects in the
XX PT p53 pathway, such as cancer of the breast, colon, kidneys, lung and
XX PT ovary.

XX PS Disclosure; Page 55-57; 82pp; English.

XX CC The invention relates to identifying a candidate p53 pathway modulating
XX CC agent in humans that is referred to in the specification as BGALT (beta-
XX CC 1,3 galactosyltransferase). Also disclosed is a method for diagnosing a
XX CC disease in a patient, by contacting a sample with a probe for B3GALT
XX CC expression, and comparing the results with a control, and determining
XX CC whether the results indicate a likelihood of disease. Methods and
XX CC compositions of the invention are useful for the diagnosis and treatment
XX CC of disorders associated with defects in the p53 pathway, such as cancer
XX CC of the breast, colon, kidneys, lung and ovary. The current sequence
XX CC represents a B3GALT nucleic acid sequence referred to in the disclosure

CC of the invention

XX SQ Sequence 3409 BP; 833 A; 842 C; 833 G; 901 T; 0 U; 0 Other;

Query Match 100.0%; Score 584; DB 10; Length 3409;
Best Local Similarity 100.0%; Pred. No. 1e-180;
Matches 584; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGACATTTATCCAGAAAGGATTTCTTAGACGCTATTACAAATCTGACCTTGAAAGACCATGA 60

Db 1430 GGGACATTTATCCAGAAAGGATTTCTTAGACGCTATTACAAATCTGACCTTGAAAGACCATGA 1489

QY 61 TGGGCATAGAAATGGGTCCATTCGCTTTTGTCTCAGCGGGCGTTTGTGATGAAAAACAGACT 120

Db 1490 TGGGCATAGAAATGGGTCCATTCGCTTTTGTCTCAGCGGGCGTTTGTGATGAAAAACAGACT 1549

QY 121 CAGACATCTTCATCAATCTGACTATCTGACTGACTGCTCTGAGAAACAGAACAA 180

Db 1550 CAGACATCTTCATCAATCTGACTATCTGACTGACTGCTCTGAGAAACAGAACAA 1609

QY 181 CCAGGTTTTTCACTGGCTTCTTGAATCTCAATGAGTTTCCCATCAGGCGAGCCATTTCAGCA 240

Db 1610 CCAGGTTTTTCACTGGCTTCTTGAATCTCAATGAGTTTCCCATCAGGCGAGCCATTTCAGCA 1669

QY 241 AGTGGTTTGTGAGTAAATCTGAATATCCGTGGGACAGGTACCCACATTTCTGCTCCGGCA 300

Db 1670 AGTGGTTTGTGAGTAAATCTGAATATCCGTGGGACAGGTACCCACATTTCTGCTCCGGCA 1729

QY 301 CCGGCTACGTGTTTCTGCGGACGTGGCGAGTCAGGTGTACAAATCTCTCCAGAGCGTCC 360

Db 1730 CCGGCTACGTGTTTCTGCGGACGTGGCGAGTCAGGTGTACAAATCTCTCCAGAGCGTCC 1789

QY 361 CATACATTTAAATCTGGAAGACGTGTTTGTGGGGCTCTGCTCGAAAGGCTGAACATCAGAT 420

Db 1790 CATACATTTAAATCTGGAAGACGTGTTTGTGGGGCTCTGCTCGAAAGGCTGAACATCAGAT 1849

QY 421 TGGAGGAGCTCCACTCCAGCCGACCTTTTTCAGGGGGCTTACGCTTCTCCGTATGCC 480

Db 1850 TGGAGGAGCTCCACTCCAGCCGACCTTTTTCAGGGGGCTTACGCTTCTCCGTATGCC 1909

QY 481 TCCTCAGGAGGATCGTGGGCTGCCACTTCATCAAGCCTCGGACTCTCTTGGACTACTGGC 540

Db 1910 TCCTCAGGAGGATCGTGGGCTGCCACTTCATCAAGCCTCGGACTCTCTTGGACTACTGGC 1969

QY 541 AGGCTCTAGAGAAATTCCTCCGGGGGAAGATTGTCCGCTGTCTGA 584

Db 1970 AGGCTCTAGAGAAATTCCTCCGGGGGAAGATTGTCCGCTGTCTGA 2013

RESULT 7

AAA93876

XX ID AAA93876 standard; DNA; 10562 BP.

XX AC AAA93876;

XX DT 15-JAN-2001 (first entry)

XX DE Human beta3Gal-T5 encoding DNA.

XX KW Beta-1,3 galactose transferase; treatment; diagnosis; cancer; human;
XX LW digestive system; beta3Gal-T5; ds.

XX OS Homo sapiens.

XX PN WO200050608-A1.

XX PD 31-AUG-2000.

XX PF 24-FEB-2000; 2000WO-JP001070.

XX PR 25-FEB-1999; 99JP-00047571.

XX KW (KYOW) KYOWA HAKKO KOGYO KK.

XX Narinatsu H, Isehiki S, Togayachi A, Sasaki K;
 XX WPI; 2000-549409/50.
 XX Beta-1,3 galactose transferase and DNA encoding it, useful for synthesis
 PT of type 1 sialyl Lewis, a carbohydrate for treatment of digestive system
 PT cancer.
 XX Claim 31; Page 103-111; 123pp; Japanese.
 XX This invention relates to a polypeptide (I) with beta-1,3 galactose
 CC transferase activity, or variants of (I) comprising amino acid additions,
 CC deletions and/or substitutions. Included in the invention is DNA encoding
 CC all or part of (I); expression vectors containing the DNA, host cells
 CC transformed by the vectors; a method for the preparation of the
 CC polypeptide by culture of the transformants or by expression in the milk
 CC of a transgenic mammal, and antibodies recognising (I). The Beta-1,3
 CC galactose transferase protein transfers galactose by beta-1,3 bonding to
 CC N-acetylglucosamine present in a non-cyclic carbohydrate chain (such as
 CC GlcNAc beta1-3Galbeta1-4Glc) to give Galbeta1-3GlcNAc. The protein and DNA
 CC encoding it are useful for the treatment and diagnosis of cancer of the
 CC digestive system. The present sequence represents a Beta3Gal-T5 encoding
 CC DNA sequence
 XX
 SQ Sequence 10562 BP; 2610 A; 2415 C; 2574 G; 2963 T; 0 U; 0 Other;
 Query Match 100.0%; Score 584; DB 3; Length 10562;
 Best Local Similarity 100.0%; Pred. No. 1.9e-180;
 Matches 584; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGACATTATCCAGAGGATTTCTAGACGCTATTACAACTCAGCCCTGAAGACCATGA 60
 DB 8593 GGGACATTATCCAGAGGATTTCTAGACGCTATTACAACTCAGCCCTGAAGACCATGA 8642
 QY 61 TGGGCATAGATGGGTCCATCGCTTTTGTCTCAGCGCGCTTTGTGATGAAACAGACT 120
 DB 8643 TGGGCATAGATGGGTCCATCGCTTTTGTCTCAGCGCGCTTTGTGATGAAACAGACT 8702
 QY 121 CAGACATGTTTCATCAATGTTGACTATCTGACTGCTTTCTGAAGAAAACAGAACAA 180
 DB 8703 CAGACATGTTTCATCAATGTTGACTATCTGACTGCTTTCTGAAGAAAACAGAACAA 8762
 QY 181 CCAGTGTTCCTACCTGGCTTTCTGAACTCAATGATGTTCCCATCAGGCGCATTCAGCA 240
 DB 8763 CCAGTGTTCCTACCTGGCTTTCTGAACTCAATGATGTTCCCATCAGGCGCATTCAGCA 8822
 QY 241 AGTGGTTTGTGATGAAATCTGAATATCCGTGGGACAGGTACCCACCATTCGTCTCCGGCA 300
 DB 8823 AGTGGTTTGTGATGAAATCTGAATATCCGTGGGACAGGTACCCACCATTCGTCTCCGGCA 8882
 QY 301 CCGGCTACGTGTTTCTGGGACGCTGGCGAGTCAAGTGTCAATGTCTTCCAGAGCGTCC 360
 DB 8883 CCGGCTACGTGTTTCTGGGACGCTGGCGAGTCAAGTGTCAATGTCTTCCAGAGCGTCC 8942
 QY 361 CATACATTAACTGGAAGACGCTGTTGGGGCTCTGCCTCGAAGGCTGAACATCAGAT 420
 DB 8943 CATACATTAACTGGAAGACGCTGTTGGGGCTCTGCCTCGAAGGCTGAACATCAGAT 9002
 QY 421 TGGAGGAGCTCCACTCCAGCGGACCTTTTTCAGGGGGCTTACGCTTCTCCGATATGCC 480
 DB 9003 TGGAGGAGCTCCACTCCAGCGGACCTTTTTCAGGGGGCTTACGCTTCTCCGATATGCC 9062
 QY 481 TCTTCAGAGGATCGTGGCCCTGCCATTCATCAAGCCTCGGACTCTCTTGGACTACTGGC 540
 DB 9063 TCTTCAGAGGATCGTGGCCCTGCCATTCATCAAGCCTCGGACTCTCTTGGACTACTGGC 9122
 QY 541 AGGCTCTAGAGATTCCTCGGGGGGAGATTGTCGCGCTCTGA 584
 DB 9123 AGGCTCTAGAGATTCCTCGGGGGGAGATTGTCGCGCTCTGA 9166

ABZ56916
 ID ABZ56916 standard; DNA; 933 BP.
 AC ABZ56916;
 XX
 DT 04-APR-2003 (first entry)
 XX
 DE B3GALT nucleic acid sequence # SEQ ID 8.
 XX
 KW B3GALT; beta-1,3 galactosyltransferase; p53; cancer; breast; colon;
 KW kidney; lung; ovary; gene; ds.
 XX Homo sapiens.
 XX WO200299044-A2.
 XX 12-DEC-2002.
 XX 02-JUN-2002; 2002WO-US017356.
 XX 05-JUN-2001; 2001US-0296076P.
 PR 10-OCT-2001; 2001US-0328605P.
 PR 15-FEB-2002; 2002US-0357253P.
 XX (EXEL-) EXELIXIS INC.
 XX
 FI Friedman L, Plozman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
 XX WPI; 2003-156849/15.
 DR
 XX
 PT Identifying p53 pathway modulating agents with B3GALT genes, useful for
 PT the diagnosis and treatment of disorders associated with defects in the
 PT p53 pathway, such as cancer of the breast, colon, kidneys, lung and
 PT ovary.
 XX
 PS Disclosure; Page 51-52; 82pp; English.
 XX
 CC The invention relates to identifying a candidate p53 pathway modulating
 CC agent in humans that is referred to in the specification as BGALT (beta-
 CC 1,3 galactosyltransferase). Also disclosed is a method for diagnosing a
 CC disease in a patient, by contacting a sample with a probe for B3GALT
 CC expression, and comparing the results with a control, and determining
 CC whether the results indicate a likelihood of disease. Methods and
 CC compositions of the invention are useful for the diagnosis and treatment
 CC of disorders associated with defects in the p53 pathway, such as cancer
 CC of the breast, colon, kidneys, lung and ovary. The current sequence
 CC represents a B3GALT nucleic acid sequence referred to in the disclosure
 CC of the invention
 XX
 SQ Sequence 933 BP; 229 A; 234 C; 240 G; 229 T; 0 U; 1 Other;
 Query Match 99.8%; Score 583; DB 10; Length 933;
 Best Local Similarity 99.8%; Pred. No. 1.1e-180;
 Matches 583; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GGGACATTATCCAGAGGATTTCTAGACGCTATTACAACTCAGCCCTGAAGACCATGA 60
 DB 350 GGGACATTATCCAGAGGATTTCTAGACGCTATTACAACTCAGCCCTGAAGACCATGA 409
 QY 61 TGGGCATAGATGGGTCCATCGCTTTTGTCTCAGGGCGGCTTTGTGATGAAACAGACT 120
 DB 410 TGGGCATAGATGGGTCCATCGCTTTTGTCTCAGGGCGGCTTTGTGATGAAACAGACT 469
 QY 121 CAGACATGTTTCATCAATGTTGACTATCTGACTGAACCTGCTTCTGAAGAAAACAGAACAA 180
 DB 470 CNGACATGTTTCATCAATGTTGACTATCTGACTGAACCTGCTTCTGAAGAAAACAGAACAA 529
 QY 181 CCAGGTTTTTCACCTGGCTTTCTTGAAGCTCAATGAGTTTCCCATCAGGCGCATTCAGCA 240
 DB 530 CCAGGTTTTTCACCTGGCTTTCTTGAAGCTCAATGAGTTTCCCATCAGGCGCATTCAGCA 589
 QY 241 AGTGGTTTGTGATGAAATCTGAATATCCGTGGGACAGGTACCCACCATTCGTCTCCGGCA 300

Db 590 AGTGTGTCAGTAAATCTGAATATCCGTGGGACAGGTACCCACCATCTGTCTCCGGCA 649
 Qy 301 CCGCTACGTGTTTCTGCGCAGCGTGGCGAGTCAAGTGTACAAATGTCTCCAGAGCGTCC 360
 Db 650 CCGCTACGTGTTTCTGCGCAGCGTGGCGAGTCAAGTGTACAAATGTCTCCAGAGCGTCC 709
 Qy 361 CATACATTTAACTGAAGACGTGTTTGTGGGCTCTGCTGCTGAAAGGCTGAACATCAGAT 420
 Db 710 CATACATTTAACTGAAGACGTGTTTGTGGGCTCTGCTGCTGAAAGGCTGAACATCAGAT 769
 Qy 421 TGGAGGAGCTCCACCTCCAGCGGACCTTTTTCAGGGGGCTTACGCTTCTCCGATGCC 480
 Db 770 TGGAGGAGCTCCACCTCCAGCGGACCTTTTTCAGGGGGCTTACGCTTCTCCGATGCC 829
 Qy 481 TCTTCAGGAGGATCGTGGGCTGCGCACTTTCATCAAGCCTCGGACTCTCTTGGACTACTGCC 540
 Db 830 TCTTCAGGAGGATCGTGGGCTGCGCACTTTCATCAAGCCTCGGACTCTCTTGGACTACTGCC 889
 Qy 541 AGGCTCTAGAGAAATTCCTGGGGGGAAGATTTGTCGCTGTCTGA 584
 Db 890 AGGCTCTAGAGAAATTCCTGGGGGGAAGATTTGTCGCTGTCTGA 933

RESULT 9

AA335710
 ID AAX35710 standard; cDNA; 1266 BP.
 XX
 AC AAX35710;

DT 09-JUL-1999 (first entry)

XX cDNA encoding a protein identified by the signal sequence trap method.

XX Signal sequence trap method; SSR method; immunisation; inhibition;
 KW infection; allergy; cancer; regulation; tissue formation; tissue repair;
 KW activin activity; inhibin activity; chemokine activity;
 KW cytokine activity; blood coagulation regulation; agonist; antagonist;
 KW metabolic disorder; hormonal disorder; immune disorder;
 KW severe combined immunodeficiency; SCID; AIDS; thrombosis; cancer; wound;
 KW ss.

XX Homo sapiens.

XX WO9918126-A1.

XX 15-APR-1999.

XX 06-OCT-1998; 98WO-JP004514.

XX 07-OCT-1997; 97JP-00274674.

XX (ONOY) ONO PHARM CO LTD.

XX Fukushima D, Shibayama S, Tada H;

XX WPI; 1999-277254/23.

XX P-PSDB; AAY02366.

XX Polypeptides identified by the signal sequence trap method from a human
 PT cDNA library.

XX Claim 4; Page 108-109; 281pp; Japanese.

XX AAX35694-X35747 represent cDNA sequences that encode novel polypeptides
 CC (AAY02358-84) which are identified from a human placental cDNA library by
 CC the signal sequence trap (SSR) method. The polypeptides have a broad
 CC range of physiological activity, including immunisation against and
 CC inhibition of infectious, allergies and cancer; regulation of tissue
 CC formation and repair; activin/inhibin activity; chemokine/cytokine
 CC activity; blood coagulation regulation; and receptor/ligand agonist or
 CC antagonist activity. The polypeptides can be used for prevention and
 CC treatment of disorders including infections by bacteria, yeasts and
 CC viruses (including HIV) and protozoa; metabolic and hormonal disorders;

CC immune disorders (including severe combined immunodeficiency (SCID) and
 CC AIDS; thrombosis; cancer; and traumatic or surgical wounds
 XX
 SQ Sequence 1266 BP; 408 A; 270 C; 242 G; 346 T; 0 U; 0 Other;

Query Match 19.6%; Score 114.4; DB 2; Length 1266;
 Best Local Similarity 51.8%; Pred. No. 2e-26;
 Matches 287; Conservative 0; Mismatches 261; Indels 6; Gaps 1;

Qy 3 GACATATCCAGAGGATTTCTAGACGTCTATTACAATCTGACCTGAAGACCATGATG 62
 Db 640 CATATAATTCACAGGAATACCTTAGATAGTACTATAATTTGACCATTTAAACACTAATG 699
 Qy 63 GGCATAGAAATGGGTCCATFCGCTTTTCTCAGCGCGGTTTGTGATGAAACACAGACTCA 122
 Db 700 GGCATGAATGGGTGCAACATACCTGTCACATATTTCCATATGTTATGAAAACTGACAGT 759
 Qy 123 GACATGTTCAATGTTGACTATCTGACTGAACCTGCTTCTGAGAGAAAACAGAACACC 182
 Db 760 GACATGTTTGTCAACACACTGAATATTTAATCAATAAAGTTACTGAAGCCAGATCTGCCTCC 819
 Qy 183 AG-----GTTTTTCACTGGCTTCTTGAAACTCAATGAGTTTCCCATCAGGCGACCATTC 236
 Db 820 AGACATAACTATTTTCACTGGTTACCTAATGAGGATATGCACCCCATCGAACAAGAT 879
 Qy 237 AGCAAGTGGTTTGTTCAGTAAATCTGAATATCCGTGGGACAGGTACCCACCATTTCTGCTCC 296
 Db 880 AGCAAGTGGTATACATGCCACACAGACCTCTACCCAGTGAGCGTTATCTCTGCTTCTTCT 939
 Qy 297 GGCACCGGCTACGTGTTTCTTGGGACGCTGGGAGTCAAGTGTACAAATGTTCTCCAGAGC 356
 Db 940 GGAACCTGGTTATGTTTCTTGGAGATCTGCAGAAAAGATTTTAAAGTTTCTTTAGGT 999
 Qy 357 GTCCCATACATTAACACTGGAAGAGCTGTTTCTGGGGCTCTGCTCGAAGGCTGAACATC 416
 Db 1000 ATCCGCGTTTGCATTTGGAAGATGATATGAGGATCTGCTTCCAGAGTTGAGATTT 1059
 Qy 417 AGATTGGAGGAGCTCCACTCCACGCGACCTTTTTTCCAGGGGGCTTACGCTTCTCCGTA 476
 Db 1060 GATCCTGTACCCCTCCCAATGAGTTTGTGTTCAATCACTGCGAGTCTCTTTATTCGAGC 1119
 Qy 477 TGCCTCTTCAGGAGGATCGTGGCTGCGCATCTCATCAAGCCTCGGACTCTCTTGGACTAC 536
 Db 1120 TGTAAATACAGCCACCTAAATTAACCTCTCATAGTTCCAGCCTAGTGAACCTGATAAATAC 1179
 Qy 537 TGGCAGGCTCTAGA 550
 Db 1180 TGGAAACCAATTTACA 1193

RESULT 10

ABZ56911
 ID ABZ56911 standard; DNA; 1268 BP.
 XX

AC ABZ56911;

DT 04-APR-2003 (first entry)

XX B3GALT nucleic acid sequence # SEQ ID 3.

XX B3GALT; beta-1,3 galactosyltransferase; p53; cancer; breast; colon;
 KW kidney; lung; ovary; gene; ds.

XX Homo sapiens.

XX WO200299044-A2.

XX 12-DEC-2002.

XX 02-JUN-2002; 2002WO-US017356.

XX 05-JUN-2001; 2001US-0296076P.

XX 10-OCT-2001; 2001US-0328605P.

Db 640 GATATAATTCAGCAGGAATATTTAGATACATATATATCTGACCAATTAACCACTAATG 699
 Qy 63 GGCATAGAAATGGGTCCATCGCTTTTGTCTCAGGGGGGTTTGTGATGAAACAGACTCA 122
 Db 700 GGTATGAATGGGTGGCACTACTGTCCACATCTCCCTATGTATGAAACGGACAGT 759
 Qy 123 GACATGTTTCAATGTTGACTATCTGACTGAATCTGCTTCTGAAAGAAAACAGAACACC 182
 Db 760 GACATGTTTGTCAACACAGAAATCTTAATACACAAAGTTACTAAAGCCAGACCTGCTCCT 819
 Qy 183 AG-----GTTTTTCACTGGCTTCTTGAATCTCAATGATTTCCCATCAGCAGCCATTC 236
 Db 820 AGACATAAATTTTACTGGCTATCTAATGAGAGGATATGACCCGAAACAGAAACAGAC 879
 Qy 237 AGCAAGTGGTTTGTCAAGTAAATCTGAATATCCGTGGGACAGGTACCCACCACTTCTGCTCC 296
 Db 880 AGTAAGTGGTATGTCACACAGACCTTTTACCCAGTGAGCGCTACCTGTCTTCTGCTCA 939
 Qy 297 GGCACCGGCTACGTTTCTTGGGACGCTGGGAGTCAAGTGTACAATGTCTTCCAAGAGC 356
 Db 940 GGAATGTTTATGTTTCTGGGATCTGGCAGAGAGATATTTAAGGTTTCTTTAGGT 999
 Qy 357 GTCCATACATTAATCTGAAGACGTGTTTGTGGGCTCTGCTCGAAAGGCTGAACATC 416
 Db 1000 ATCCGCTGTTTGCACCTTGGAAAGATGATATGTAGGGATCTGTCTTGGCCAAAGTTGAGAGTT 1059
 Qy 417 AGATTGGAGGAGCTCCACTCCCGAGCGACCTTTTTCAGGGGCTTACGGCTTCTCCGTA 476
 Db 1060 GATCCTGTGCCCCCTCCCAATGAGTTCCTCATCAGTTCCAGCCTAGTGAACCTGATAAATAC 1179

RESULT 12
 ABZ56913 ID ABZ56913 standard; DNA; 1269 BP.
 XX AC ABZ56913;
 XX AC
 DT 04-APR-2003 (first entry)
 XX B3GALT nucleic acid sequence # SEQ ID 5.
 DE B3GALT; beta-1,3 galactosyltransferase; p53; cancer; breast; colon;
 KW kidney; lung; ovary; gene; ds.
 XX Homo sapiens.
 OS
 XX WO200299044-A2.
 XX PN
 PD 12-DEC-2002.
 XX
 XX 02-JUN-2002; 2002WO-US017356.
 XX
 XX 05-JUN-2001; 2001US-0296076P.
 PR 10-OCT-2001; 2001US-0328605P.
 PR 15-FEB-2002; 2002US-0357253P.
 XX
 XX (EXEL-) EXELIXIS INC.
 FA
 XX Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
 PI
 XX WPI; 2003-156849/15.
 DR
 XX Identifying p53 pathway modulating agents with B3GALT genes, useful for
 PT the diagnosis and treatment of disorders associated with defects in the

PT p53 pathway, such as cancer of the breast, colon, kidneys, lung and
 PT ovary.
 XX
 XX Example 5; Page 45; 82pp; English.
 CC The invention relates to identifying a candidate p53 pathway modulating
 CC agent in humans that is referred to in the specification as B3GALT (beta-
 CC 1,3 galactosyltransferase). Also disclosed is a method for diagnosing a
 CC disease in a patient, by contacting a sample with a probe for B3GALT
 CC expression, and comparing the results with a control, and determining
 CC whether the results indicate a likelihood of disease. Methods and
 CC compositions of the invention are useful for the diagnosis and treatment
 CC of disorders associated with defects in the p53 pathway, such as cancer
 CC of the breast, colon, kidneys, lung and ovary. The current sequence
 CC represents a B3GALT nucleic acid sequence referred to in an example from
 CC the invention
 XX
 SQ Sequence 1269 BP; 409 A; 269 C; 243 G; 348 T; 0 U; 0 Other;
 Query Match 19.6%; Score 114.4; DB 10; Length 1269;
 Best Local Similarity 51.8%; Pred. No. 2e-26;
 Matches 287; Conservative 0; Mismatches 261; Indels 6; Gaps 1;
 Qy 3 GACATATTCAGAGAGATTTCCTAGACGTCTATTACATCTGACCTGAAGCAATGATG 62
 Db 640 GATATAATTCACAGGAATACTTAGATACGTAATATTTGACCAATTAACCACTAATG 699
 Qy 63 GGCATAGAAATGGGTCCATCGCTTTTGTCTCAGGGGGGTTTGTGATGAAACAGACTCA 122
 Db 700 GGCATAGAAATGGGTGGCAACATCTGTCCACATATTTCCATATGTTATGAAACTGACAGT 759
 Qy 123 GACATGTTTCAATGTTGACTATCTGACTGAACTGCTTCTGAAAGAAAACAGAAACACC 182
 Db 760 GACATGTTTGCACACTGATATTTAATCAATGTTTACTGAAGCAGATCTGCTCCTCC 819
 Qy 183 AG-----GTTTTTCACTGGCTTCTGAAACTCAATGATTTCCCATCAGCAGCCATTC 236
 Db 820 AGACATAAATTTTCACTGGTTTACCTAATGCGAGGATATGCACCAATCGAAACAAAGAT 879
 Qy 237 AGCAAGTGGTTTGTCAAGTAAATCTGAATATCCGTGGGACAGTACCCACCACTTCTGCTCC 296
 Db 880 AGCAAGTGGTATGATGCCACAGACTCTTACCAAGTGAGCGTTATCTGTCTTCTGTTCT 939
 Qy 297 GGCACCGGCTACGTTGTTTCTTGGGACGCTGCGAGTCAAGTGTACAATGTCTTCCAAGAGC 356
 Db 940 GGAATGTTTATGTTTCTTGGAGATCTGCAGAAAGAGATTTTAAAGTTTCTTTAGGT 999
 Qy 357 GTCCATACATTAATCTGAAGACGTGTTTGTGGGCTCTGCTCGAAAGGCTGAACATC 416
 Db 1000 ATCCGCGGTTTGCACCTTGGAAAGATGATATGTAGGGATCTGTCTTGGCCAAAGTTGAGAAAT 1059
 Qy 417 AGATTGGAGGAGCTCCACTCCCGAGCGACCTTTTTCAGGGGCTTACGGCTTCTCCGTA 476
 Db 1060 GATCCTGTACCCCTCCCAATGAGTGTGTTGTTCAATCACTGCGGAGTCTTTATTCGAGC 1119
 Qy 477 TGCCTCTTTCAGGAGGATCGTGGCTGCGACCTTCAATCAAGCCTCGGACTCTCTTGGACTAC 536
 Db 1120 TGTAATACAGCCACCTAATTTACCTCTCATCAGTTCCAGCCTAGTGAACCTGATAAATAC 1179
 Qy 537 TGGCAGGCTCTAGA 550
 Db 1180 TGGAACCAATTTACA 1193
 RESULT 13
 ADM66890 ID ADM66890 standard; DNA; 1473 BP.
 XX AC ADM66890;
 XX AC
 DT 03-JUN-2004 (first entry)
 DE Murine adipocyte specific beta 1,3 galactosyltransferase DNA SeqID 23.

XX murine; mouse; adipocyte specific; gene; ds; adipose tissue;
 KW anti-obesity; high mobility group I-C protein; HMGI-C; obesity; leptin;
 KW ob; diabetes; adipogenesis; hypertension; cardiovascular disease;
 KW anorectic; antidiabetic; hypotensive; beta 1,3 galactosyltransferase.
 XX
 OS Mus musculus.
 XX
 PN W0204011618-A2.
 XX
 PD 05-FEB-2004.
 XX
 PP 29-JUL-2003; 2003WO-US023684.
 XX
 PR 29-JUL-2002; 2002US-0398785P.
 PR 12-JUN-2003; 2003US-0478206P.
 XX
 PA (HMGH-) HMGNE INC.
 XX
 PI Chada K, Chouinard R, Achar H, Sayed AMD;
 XX
 DR WPI; 2004-143846/14.
 DR P-PSDB; ADM67170.
 XX
 PT Identifying adipocyte specific genes, useful for treating obesity or
 PT diabetes, and for identifying drug targets, by differential gene
 PT expression analysis between adipose tissue or stromal vascular tissue of
 PT mice of different genotypes.
 XX
 PS Claim 11; SEQ ID NO 23; 91pp; English.
 XX
 CC This invention relates to a novel method for identifying genes that are
 CC over-expressed in adipose tissue and as such it provides targets for anti-
 CC -obesity pharmaceutical compositions. Specifically, it refers to a high
 CC mobility group I-C protein (HMGI-C) that is associated with obesity and
 CC is epistatic to leptin, furthermore, it refers to the ob gene where an
 CC autosomal recessive trait is linked to obesity and diabetes. The present
 CC invention describes performing differential gene expression analysis
 CC between the white adipose tissue (WAT) or stromal vascular tissue (SVT)
 CC of any two different mice selected from a group consisting of wild-type,
 CC HMGI-C -/-, ob/ob, or HMGI-C -/- ob/ob genotype mice. Accordingly, using
 CC this method novel nucleotides and the encoded proteins thereof were
 CC identified that are adipocyte specific, and as such can be used for
 CC preventing adipogenesis, diagnosing and treating diabetes, obesity,
 CC hypertension and cardiovascular disease, as well as screening for
 CC compounds that can modulate or prevent adipogenesis and treat diabetes or
 CC obesity. These compositions exhibit anorectic, antidiabetic and
 CC hypotensive activities. This polynucleotide sequence is a murine
 CC adipocyte specific DNA sequence of the invention.
 XX
 SQ Sequence 1473 BP; 470 A; 315 C; 288 G; 400 T; 0 U; 0 Other;
 Query Match
 Best Local Similarity 19.6%; Score 114.4; DB 12; Length 1473;
 Matches 287; Conservative 0; Mismatches 261; Indels 6; Gaps 1;
 QY 3 GACATATCCAGAAGATTCTAGACGTCTATTACAACTGACCTGAAGACCATGATG 62
 Db 702 GATATAATTCAGCAGGATATTAGATACATACATATATCTGACCATTTAAACACTAATG 761
 QY 53 GGCATAGATGGGTCCATCGCTTTTGTCTCAGGGGGGTTTGTGATGAAACAGACTCA 122
 Db 712 GGTATGAATCGGGTTGCAACACTACTGTCCACATCTCCCTATGTTATGAAACCGACAGT 821
 QY 183 GACATGTTTCATCAATGTTGACTATCTGACTGAACCTCTTGAAGAAAAACAGAACACC 182
 Db 812 GACATGTTTGTCAACACAGATACATCTATACAGATTACTAAAGCCAGACTGCCTCT 881
 QY 183 AG-----GTTTTTCACGTGGCTTCTTGAACACTCAATGATGTTTCCCATCAGGACGCATTC 236
 Db 812 AGACATAACTATTTTACTGGCTTCTTAAATGAGAGGATATGCACCGAACAGAAACAAAGAC 941
 QY 217 AGCAAGTGGTTGTTCAGTAAATCTGAATATCCGTGGGACAGGTACCCACCATTCTGCTCC 296

Db 942 AGTAAGTGGTACATGCCACGACCTTTTACCCAGTGAGCGCTACCCCTGTCTCTGCTCA 1001
 QY 297 GGCACCGGCTACGTGTTTCTGGGAGCTGGGAGTCAGGTGTACAAATGTCTCCAAAGC 356
 Db 1002 GGAACCTGTTATGTTTCTGGGGATCTGGCAGAGAGATATTTAAGGTTTCTTTAGGT 1061
 QY 357 GTCCCATACATTAACTGGAAGACGTGTTTGTGGGCTCTGCCCTCGAAAGGCTGAACATC 416
 Db 1062 ATCCGTCGTTTGCACCTTGGAGATGTATATGTAGGGAATCTGCTTGCCTTGAAGTT 1121
 QY 417 AGATTGGAGGAGCTCCACTCCAGCCGACCTTTTTCAGGGGGCTTACGCTTCTCCGTA 476
 Db 1122 GATCCTGTGCCCCCTCCCAATGAGTTGCTGTTCAATCACTGGCGAGTTCTTATTCAAGC 1181
 QY 477 TGCCTCTTTCAGGAGATGCTGGGCTGCTCACTTCAATCAAGGCTCGGACTCTCTTGGACTAC 536
 Db 1182 TGTAAATACAGCCACCTAAATTAACCTCTCATCAGTTCAGGCTAGTGAACCTGATAAATAC 1241
 QY 537 TGGCAGGCTCTAGA 550
 Db 1242 TGGAACCAATTTACA 1255
 RESULT 14
 AAX35711
 ID AAX35711 standard; cDNA; 2420 BP.
 XX
 AC AAX35711;
 XX
 DT 09-JUL-1999 (first entry)
 XX
 DE cDNA encoding a protein identified by the signal sequence trap method.
 KW Signal sequence trap method; SST method; immunisation; inhibition;
 KW infection; allergy; cancer; regulation; tissue formation; tissue repair;
 KW activin activity; inhibin activity; chemokine activity;
 KW cytokine activity; blood coagulation regulation; agonist; antagonist;
 KW metabolic disorder; hormonal disorder; immune disorder;
 KW severe combined immunodeficiency; SCID; AIDS; thrombosis; cancer; wound;
 KW ss.
 XX
 OS Homo sapiens.
 XX
 PN W09918126-A1.
 XX
 PD 15-APR-1999.
 XX
 PF 06-OCT-1998; 98WO-JP004514.
 XX
 PR 07-OCT-1997; 97JP-00274674.
 XX
 PA (ONOY) ONO PHARM CO LTD.
 XX
 PI Fukushima D, Shibayama S, Tada H;
 XX
 DR WPI; 1999-277254/23.
 XX
 P-PSDB; AAY02366.
 XX
 PT Polypeptides identified by the signal sequence trap method from a human
 XX cDNA library.
 PS Claim 5; Page 110-113; 281pp; Japanese.
 XX
 CC AAX35694-X35747 represent cDNA sequences that encode novel polypeptides
 CC (AAY02358-84) which are identified from a human placental cDNA library by
 CC the signal sequence trap (SST) method. The polypeptides have a broad
 CC range of physiological activity, including immunisation against and
 CC inhibition of infections, allergies and cancer; regulation of tissue
 CC formation and repair; activin/inhibin activity; chemokine/cytokine
 CC activity; blood coagulation regulation; and receptor/ligand agonist or
 CC antagonist activity. The polypeptides can be used for prevention and
 CC treatment of disorders including infections by bacteria, yeasts and

CC	viruses (including HIV) and protozoa; metabolic and hormonal disorders;
CC	immune disorders (including severe combined immunodeficiency (SCID) and
CC	AIDS; thrombosis; cancer; and traumatic or surgical wounds
XX	
SQ	Sequence 2420 BP; 853 A; 458 C; 451 G; 658 T; 0 U; 0 Other;
Query Match	19.6%; Score 114.4; DB 2; Length 2420;
Best Local Similarity	51.8%; Pred. No. 2.9e-26;
Matches 287; Conservative	0; Mismatches 261; Indels 6; Gaps 1;
QY	3 GACATTATCCAGAGGATTTCTTAGAGCTCTATTACAACTCGACCTCGAAGACCATCATG 62
DB	1337 GATATAATTCAACAGGAATACCTTAGATACGTACTATAATTTGACCATTAATAAACAATAAG 1396
QY	63 GGCATGAATGGGTCCATCGCTTTGTCTCCAGCGCGGTTTGTGATGAAAAACAGACTCA 122
DB	1397 GGCATGAATGGGTGCAACATACGTGCCAATATCCATATGTTATGAAAACTGACAGT 1456
QY	123 GACATGTTTCATCAATGTTGACTATCTGACTGAACATGCTTCTGAAAGAAAAACAGAAACC 189
DB	1457 GACATGTTTGTCAACACTGAATATTTAATCAATAAGTTACTGAAAGCAGATCTGCCTCCC 1516
QY	183 AG-----GTTTTTCACTGGCTTCTGAAACTCAATCAGTTTCCCATCAGCGACCAATC 236
DB	1517 AGACATAACTATTTCATCTGGTTACTTAATGCGAGGATATGCACCCATCGAAACAAAGAT 1576
QY	237 AGCAAGTGGTTTTGTTCAGTAAATCTGAAATATCCGTGGGACAGGTACCCACCAATTCGTCTCC 296
DB	1577 AGCAAGTGGTACATGCCACCAGACCTCTACCCAAAGTCAGCGTTATCTGTCTCTGTCT 1636
QY	297 GGCACCGGCTACGTTTCTTCGCGAGCGTGGCGAGTCAGGTGTACAATGTCCTCAAGAGC 356
DB	1637 GGAACCTGTTATGTTTTTCTTCGGAGATCTGGCAGAAAAGATTTTAAAGTTTCTTTAGGT 1696
QY	357 GTCCCATACATTAACCTGGAAGACGCTTTGTGCGGCTCTGCCTCGAAAGCGCTGAACATC 416
DB	1697 ATCCGCGTTTGCACTTTGGAAGATGTATATGTAGGATCTGTCTTGCCAAAGTTGAGAATT 1756
QY	417 AGATTGAGGAGTGTCACCTCCACGCCACCTTTTTTCCAGGGGCTTACGCTTCTCCGTA 476
DB	1757 GATCCTGTACCCCTCCCAATGAGTTGTGTTCAATCACTGGCGAGTCTCTTATTCGAGC 1816
QY	477 TGCTCTTTCAGAGGATCGTGGCGTCCCATTTCAATCAAGCGCTCGGACTCTCTTTGGACTAC 536
DB	1817 TGTAAATACAGCCACCTAAATTACCTCTCATCAGTTCACAGCTAGTGAACCTGATAAAATAC 1876
QY	537 TGGCAGGCTCTAGA 550
DB	1877 TGAACCAATTTACA 1890
RESULT 15	
ADM66891	ADM66891 standard; DNA; 3054 BP.
XX	AC
XX	ADM66891;
XX	
DT	03-JUN-2004 (first entry)
XX	
DE	Human adipocyte specific beta 1.3 galactosyltransferase DNA SeqID 24.
XX	
KW	human; adipocyte specific; gene; ds; adipose tissue; anti-obesity;
KW	high mobility group I-C protein; HMG1-C; obesity; leptin; ob; diabetes;
KW	adipogenesis; hypertension; cardiovascular disease; anorectic;
KW	antidiabetic; hypotensive; beta 1,3 galactosyltransferase.
XX	
OS	Homo sapiens.
XX	
PN	W02004011618-A2.
XX	
PD	05-FEB-2004.
XX	
PF	29-JUL-2003; 2003WO-US023684.

Db 1949 GATCCTGTACCCCTCCCAATGAGTTTGTGTTCAATCACTGGCGAGTCTCTTATTGAGC 2008
Qy 477 TGCCTCTTCAGGAGGATCGTGGCCCTGCCACTTCATCAGCCCTCGGACTCTCTTGGACTAC 536
Db 2009 TGTAATACAGCCACCTTAATTACCTCTCATCAGTTCCAGCCCTAGTGAAGTGAATAAATAC 2068
Qy 537 TGGCAGGCTCTAGA 550
Db 2069 TGGACCAATTACA 2082

Search completed: September 22, 2005, 12:38:54
Job time : 351.351 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2005, 12:09:12 ; Search time 2575.06 Seconds
(without alignments)
8632.613 Million cell updates/sec

Title: US-10-777-828-8_COPY_428_1011

Perfect score: 584

Sequence: 1 gggacattatccagaagat.....aagattgcgcctgtctga 584

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 1903213700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

BST:*

- 1: gb_est1:*
- 2: gb_est2:*
- 3: gb_hic:*
- 4: gb_est3:*
- 5: gb_est4:*
- 6: gb_est5:*
- 7: gb_est6:*
- 8: gb_ges1:*
- 9: gb_ges2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	584	100.0	933	9	AY419460 Homo sapi
2	419.4	71.8	780	9	AY419461 Pan trogl
3	369.6	63.3	434	2	BF748181 RC3-BN041
4	368	63.0	5069	3	BC047347 Mus muscu
5	363.2	62.2	927	9	AY419462 Mus muscu
6	282.4	48.4	766	9	CE132434 tigr-ges-
7	184.2	31.5	629	7	CR564113 CR564113
8	173	29.6	493	1	AJ507051 AJ507051
9	153.8	26.3	492	5	BP761359 BP761359
10	148.2	25.4	626	5	EX302370 EX302370
11	131	22.4	901	4	BJ740020 BJ740020
12	122.8	21.0	981	9	AY401366 Mus muscu
13	122.8	21.0	2705	3	AK035215 Mus muscu
14	121.2	20.8	685	5	BUI133295 BUI133295
15	121.2	20.8	842	5	BUI30312 BUI30312
16	120	20.5	447	6	CB546569 CB546569
17	120	20.5	1101	9	CNS05BT1 CNS05BT1
18	119	20.4	1903	3	AK088407 AK088407
19	117.8	20.2	879	5	BQ95238 BQ95238
20	117.4	20.1	996	9	AY402383 AY402383
21	117.4	20.1	1514	3	AK003937 AK003937
22	116	19.9	596	4	BU501651 BU501651
23	116	19.9	644	6	CA057816 CA057816
24	114.6	19.6	697	6	CB233493 CB233493

25	114.4	19.6	582	5	BP374190 BP374190
26	114.4	19.6	1238	9	AY420414 AY420414
27	114	19.5	444	4	BM403751 BM403751
c	28	114	574	2	BE652985 BE652985
29	113.4	19.4	810	7	CF749272 CF749272
30	112.8	19.3	682	6	BY704097 BY704097
31	112.8	19.3	787	7	CK475539 CK475539
32	112.8	19.3	994	7	CF585163 CF585163
33	112.8	19.3	1238	9	AY420416 AY420416
34	112.8	19.3	2681	3	AK036141 AK036141
35	112.8	19.3	2690	3	AK083168 AK083168
36	112.8	19.3	4427	3	AK034371 AK034371
37	111.6	19.1	981	9	AY401365 AY401365
c	38	111.2	632	4	BJ632719 BJ632719
39	110.6	18.9	981	9	AY401364 AY401364
40	109.6	18.8	729	7	CF113712 CF113712
c	41	108.6	632	9	CC554157 CC554157
42	108.4	18.6	965	9	CNS05IB2 CNS05IB2
43	108.2	18.5	831	7	C0398909 C0398909
44	107.8	18.5	1020	7	CN646142 CN646142
45	106.2	18.2	712	1	AL042887 AL042887

ALIGNMENTS

RESULT 1
LOCUS AY419460 933 bp DNA linear GSS 17-DEC-2003
DEFINITION Homo sapiens B3GALT5 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION AY419460
VERSION AY419460.1 GI:39775417
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 933)
AUTHORS Clark,A.G., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 933)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT These sequences were made by sequencing genomic exons and ordering them based on alignment.
FEATURES
source
Location/Qualifiers
1..933
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>933
/gene="B3GALT5"
/locus_tag="HCN6898"

Query Match 100.0%; Score 584; DB 9; Length 933;
Best Local Similarity 100.0%; Pred. No. 6.1e-165;
Matches 584; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGACATTATCCAGAAGATTTCCTAGACGTCCTATTACATCTGACCTCGAAGACCATGA 60
|||||

```
Db 350 GGGACATTATCCAGAAGGATTTCCTAGACGTCTATTACAACTGACCTGAAGACCATGA 409
Qy 51 TGGGCATAGATGGGTCCATCGCTTTTGTCTCAGCGCGGCTTTGTGATGAACACAGACT 120
Db 410 TGGGCATAGATGGGTCCATCGCTTTTGTCTCAGCGCGGCTTTGTGATGAACACAGACT 469
Qy 151 CAGACATGTTTCATCAATCAATGTTGACTATCTGACTGAACTCTTCTGAAGAAAAACAGAACAA 180
Db 470 CAGACATGTTTCATCAATGTTGACTATCTGACTGAACTCTTCTGAAGAAAAACAGAACAA 529
Qy 181 CCAGGTTTTTCACTGGCTTCTTGAAGAACTCAATGAGTTTTCCCATCAGGCAGCCATTGACGA 240
Db 530 CCAGGTTTTTCACTGGCTTCTTGAAGAACTCAATGAGTTTTCCCATCAGGCAGCCATTGACGA 589
Qy 241 AGTGGTTTGTGAGTAATCTGAATATCCGTGGGACAGGTACCCACCATCTGCTCCGGCA 300
Db 590 AGTGGTTTGTGAGTAATCTGAATATCCGTGGGACAGGTACCCACCATCTGCTCCGGCA 649
Qy 301 CCGGCTACGTGTTTCTGGCGAGCTGGCGAGTCAAGTGTGACAACTCTCCAAAGAGCGTCC 360
Db 650 CCGGCTACGTGTTTCTGGCGAGCTGGCGAGTCAAGTGTGACAACTCTCCAAAGAGCGTCC 709
Qy 361 CATACATTAACTGGAAGACGTGTTTGTGGGCTCTGCTCGAAAGGCTGAACATCAGAT 420
Db 710 CATACATTAACTGGAAGACGTGTTTGTGGGCTCTGCTCGAAAGGCTGAACATCAGAT 769
Qy 421 TGGAGAGCTCAGTCCACAGCGACTTTTTCAGGGGGCTTACGCTTCTCCGTATGCC 480
Db 770 TGGAGAGCTCAGTCCACAGCGACTTTTTCAGGGGGCTTACGCTTCTCCGTATGCC 829
Qy 481 TCTTCAGAGAGATCGTGGCTGCCACTTTCATCAAGCTCGGACTCTCTTGGACTACTGCG 540
Db 830 TCTTCAGAGAGATCGTGGCTGCCACTTTCATCAAGCTCGGACTCTCTTGGACTACTGCG 889
Qy 541 AGGCTCTAGAGAAATCCCGGGGGGAAGATTGTCCGCTGTCTGA 584
Db 890 AGGCTCTAGAGAAATCCCGGGGGGAAGATTGTCCGCTGTCTGA 933

RESULT 2
AY419461 780 bp DNA linear GSS 17-DEC-2003
LOCUS Pan troglodytes B3GALT5 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY419461
VERSION AY419461.1 GI:39775418
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1 (bases 1 to 780)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarimal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 780)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarimal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
These sequences were made by sequencing genomic exons and ordering
them based on alignment.
LOCATION/Qualifiers
1. .780
/organism="Pan troglodytes"
```

```
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>780
/gene="B3GALT5"
/locus_tag="HCM6898"

ORIGIN
Query Match 71.8%; Score 419.4; DB 9; Length 780;
Best Local Similarity 98.1%; Pred. No. 3.3e-115; Indels 0; Gaps 0;
Matches 423; Conservative 0; Mismatches 8;

Qy 1 GGGACATTATCCAGAAGGATTTCCTAGACGTCTATTACAACTGACCTGAAGACCATGA 60
Db 350 GGGACATCATCCAGAAGGATTTCCTGACGCTCTATTACAACTGACCTGAAGACCATGA 409
Qy 61 TGGGCATAGATGGGTCCATCGCTTTTGTCTCAGCGCGGCTTTGTGATGAACACAGACT 120
Db 410 TGGGCATAGATGGGTCCATCGCTTTTGTCTCAGCGCGGCTTTGTGATGAACACAGACT 469
Qy 121 CAGACATGTTTCATCAATGTTGACTATCTGACTGAACTCTTCTGAAGAAAAACAGAACAA 180
Db 470 CAGACATGTTTCATCAATGTTGACTATCTGACTGAACTCTTCTGAAGAAAAACAGAACAA 529
Qy 181 CCAGGTTTTTCACTGGCTTCTTGAAGAACTCAATGAGTTTTCCCATCAGGCAGCCATTGACGA 240
Db 530 CCAGGTTTTTCACTGGCTTCTTGAAGAACTCAATGAGTTTTCCCATCAGGCAGCCATTGACGA 589
Qy 241 AGTGGTTTGTGAGTAATCTGAATATCCGTGGGACAGGTACCCACCATCTGCTCCGGCA 300
Db 590 AGTGGTTTGTGAGTAATCTGAATATCCGTGGGACAGGTACCCACCATCTGCTCCGGCA 649
Qy 301 CCGGCTACGTGTTTCTGGCGAGCTGGCGAGTCAAGTGTGACAACTCTCCAAAGAGCGTCC 360
Db 650 CCGGCTACGTGTTTCTGGCGAGCTGGCGAGTCAAGTGTGACAACTCTCCAAAGAGCGTCC 709
Qy 361 CATACATTAACTGGAAGACGTGTTTGTGGGCTCTGCTCGAAAGGCTGAACATCAGAT 420
Db 710 CATACATTAACTGGAAGACGTGTTTGTGGGCTCTGCTCGAAAGGCTGAACATCAGAT 769
Qy 421 TGGAGAGGCTC 431
Db 770 TGGAGGAGCTC 780

RESULT 3
BF748181/c
LOCUS BF748181
DEFINITION RC3-BN0411-021000-021-c04 BN0411 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF748181
VERSION BF748181.1 GI:12074857
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 434)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Ngai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Mateukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
PUBMED
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
```

Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC3&t2=RC3-BN0411-021000-021-c04&t3=2000-10-02&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 24
 High quality sequence stop: 432.

FEATURES
 Location/Qualifiers
 1..434
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="BN0411"
 /note="Organ: breast normal; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN
 Query Match 63.3%; Score 369.6; DB 2; Length 434;
 Best Local Similarity 98.9%; Pred. No. 3.2e-100;
 Matches 372; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 209 CAATGAGTTTCCCATCAGCAGCCATTTCAGCAAGTGGTTTGTCAAGTAATCTGAATATCC 268
 DB 434 CAATGAGTTTCCCATCAGCAGCCATTTCAGCAAGTGGTTTGTCAAGTAATCTGAATATCC 375
 QY 269 GTGGACAGGTACCCACCACTTCGTCCGGCACCGGTACGTTTTCGTGGCAGGTGCG 328
 DB 374 GTGGACAGGTACCCACCACTTCGTCCGGCACCGGTACGTTTTCGTGGCAGGTGCG 315
 QY 329 GAGTCAGGTGTCAATGTCTTCAAGAGCGTCCCATATCAATTAACATGGAAGAGCTGTTGT 388
 DB 314 GAGTCAGGTGTCAATGTCTTCAAGAGCGTCCCATATCAATTAACATGGAAGAGCTGTTGT 255
 QY 389 GGGGCTTCCTCGAAGCTGAACATCATGATGGAGAGTCCACTCCAGCCGACCTT 448
 DB 254 GGGGCTTCCTCGAAGCTGAACATCATGATGGAGAGTCCACTCCAGCCGACCTT 195
 QY 449 TTTTCCAGGGGCTTACGCTTCTCGTATGCTCTTTCAGAGAGTGTGGCTGCGCCTT 508
 DB 194 TTTTCCAGGGGCTTACGCTTCTCGTATGCTCTTTCAGAGAGTGTGGCTGCGCCTT 135
 QY 509 CATCAAGCTCGGACTCTCTTGAGTACTTGGCAGGCTCTAGAGAATTCCTCCGGGGGAAGA 568
 DB 134 CATCAAGCTCGGACTCTCTTGAGTACTTGGCAGGCTCTAGAGAATTCCTCCGGGGGAAGA 75
 QY 569 TTGTCGCGCTGTCTGA 594
 DB 74 TTGTCGCGCTGTCTGA 59

RESULT 4
 BC047347
 LOCUS
 DEFINITION Mus musculus UDP-Gal:betaGalCNAC beta 1,3-galactosyltransferase, polypeptide 5, mRNA (cDNA clone IMAGE:5365404).
 ACCESSION BC047347
 VERSION BC047347.1 GI:28611171
 KEYWORDS HTC.
 SOURCE Mus musculus
 ORGANISM Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 5069)
 STRAUSBERG,R.D., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,

Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.P., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.P., Casavant,T.L., Schaeetz,T.E., Brownstein,M.J., Uedin,T.B., Toshiyuki,S., Carninci,P., Frange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahy,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalley,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 5069)
 Strausberg,R.
 Direct Submission
 Submitted (28-FEB-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: http://mgc.nci.nih.gov
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: The Cepko Laboratory
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
 Center code: BCM-HGSC
 Web site: http://www.hgsc.bcm.tmc.edu/cdna/
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulsegod, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Series: IRAP Plate: 53 Row: j Column: 9
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 22122432
 This clone has the following problem: no 5' EST match.

FEATURES
 Location/Qualifiers
 1..5069
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:5365404"
 /tissue_type="Eye, retina, mouse strain C57Bl/6"
 /clone_lib="NIH MGC_94"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"

ORIGIN
 Query Match 63.0%; Score 368; DB 3; Length 5069;
 Best Local Similarity 76.9%; Pred. No. 2.1e-99;
 Matches 449; Conservative 0; Mismatches 135; Indels 0; Gaps 0;

QY 1 GGGACATTTATCAGAGGATTTCTTAGAGCTTATTACATCTGACCTGAGACCATGA 60
 DB 935 GCGACATTATCAGAGGACTTCAAGGATGCTTCAACCTGACCTGAGACCATGA 994
 QY 61 TGGGATAGATGGTCCATCGCTTTCTCTCAGCGCGCTTCTGTGATGAACAGACT 120
 DB 995 TGGGATAGATGGTCTTACCACCTTTTGTCTCTCAGACAGCTTACGTGATGAACAGGACT 1054
 QY 121 CAGACATGTTCAATGTTGACTATCTGACTGAACTGCTTCTGAAGAAAAACAGAACAA 180

```

1055 CTGACATGTTCTGTAATGTTGGCTATCTGACGGAACTGCTGTAAGAAACAAACAAAGA 1114
181 CGAGTTTTCACCTGGCTCTTTGAAACTCAATGAGTTTCCCATCAGGAGCAATTCAGCA 240
1135 CCAGGTTCTTTCACAGGCTACATATAAGCCCAAGCTTTCCCATCCGGCAGAAAGTTCAACA 1174
241 AGTGGTTTGTCTAGTAATCTGAATACGTTGGGACAGGTACCCACCATCTGCTCCGGCA 300
1175 AGTGGTTTGTGAGTAAGTTTGAATATCCCTGGGACAGGTACCCACCTTTTGTCTCCGTA 1234
301 CCGGCTAGCTGTTTCTGGCGACGTTGGCGAGTCAGGTGTACAATCTCTCCAGAGCGTCC 360
1235 CTGGTTATGTTCTTTCCAGCGAGTGGCGATCCAGTATACAAGTCTCAGAGAGCGTTC 1294
361 CATACATTAACATGGAAGACGTTGTTTGGGGCTCTGCTCGAAAGGCTGAACATCAGAT 420
1295 GCTTCATCAAGCTGAGGATGTTTGTGGGCTCTGCTTGGCCAAAGCTAAAGATCCGGC 1354
421 TGGAGGAGCTCCACTCCAGCGGACCTTTTTCAGGGGGCTTACGCTTCTCGTATGCC 480
1355 CGGAGGAGCTGCACACCAACAGACCTTCTTCCCTGGCGGTTTACGCTTTTCCGGTGCC 1414
481 TCTTCAGGAGGATCGTGGCTGCGCACTTCATCAAGCCTCGGACTCTCTTGGACTACTGCC 540
1415 GCTTTCAGAAATTTGGGATGCGCATTTTATGAAGCCCGAGGACCTGCTCACTACTGCC 1474
541 AGGCTCTAGAGAAATCCCGGGGGAGATGTTCCGCTGTCTGA 584
1475 AAGCACTGGAGAACTCGAAAGAACAGGACTGTCTGCTGTCTGA 1518

RESULT 5
AY419462
Mus musculus B3GALT5 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
AY419462.1 GI:39775419
GSS.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 927)
Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarival, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 927)
Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarival, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
These sequences were made by sequencing genomic exons and ordering
them based on alignment.
Location/Qualifiers
1..927
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>927
/gene="B3GALT5"
/locus_tag="HMC6898"

```

```

Query Match 62.2%; Score 363.2; DB 9; Length 927;
Best Local Similarity 76.4%; Pred. No. 3.4e-98;
Matches 446; Conservative 0; Mismatches 138; Indels 0; Gaps 0;
QY 1 GGGCATTATCCAGAAGGATTTCTAGACGCTCTATTACAATCTGACCTGAGACCATGA 60
Db 344 GCGACATTATCCAGAAGGACTTCAAGGATGCTTCAACTGACCTGAGAAACCATGA 403
QY 61 TGGGCATAGATGGTCCATCGCTTTTGTCTCAGGGGGGTTTGTGATGAAGAACAGACT 120
Db 404 TGGGTATGGATGGGCTCTACCACTTTTGTCTCAGACAGCTTACGTGATGAAGAACGGACT 463
QY 121 CAGACATGTTTCATCAATGTTTGAACCTCAATAGTCTTCTCAAGAAAAACAGAAACA 180
Db 464 CTGACATGTTTGTGAATGTTGGCTATCTGACGGAACCTGCTCTAAGAAAAACAGAAACA 523
QY 181 CCAGGTTTTTCATCGGCTTTTGAACCTCAATAGTCTTCCCATCAGGAGCCATTCAGCA 240
Db 524 CCAGGTTCTTCAAGGCTTACATAAAGCCCAAGCTTTTCCCATCCGGCAGAAAGTTCAACA 583
QY 241 AGTGGTTTGTGAGTAATCTGAATATCGTGGGACAGTACCCACCATCTTGTCTCCGGCA 300
Db 584 AGTGGTTTGTGAGTAATCTGAATATCGTGGGACAGTACCCACCATCTTGTCTCTGTA 643
QY 301 CCGGCTAGCTGTTTCTGGCGACGTTGGCGAGTCAGGTGTACAATGTTCTCCAGAGCGTCC 360
Db 644 CTGGTTATGTTCTTTCCAGCGAGCTGGCGATCCAGTATACAAGTCTCAGAGAGCGTTC 703
QY 361 CATACATTAAACTGGAAGACGTTGTTGGGGCTCTGCTCGAAAGGCTGAACATCAGAT 420
Db 704 GCTTCATCAAGCTGGAGGATGTTGTTGGGCTCTGCTTGGCCAAAGCTAAAGATCCGGC 763
QY 421 TGGAGGAGCTCCACTCCAGCGGACCTTTTTCAGGGGGCTTACGCTTCTCCGATGCC 480
Db 764 CGGAGGAGCTGCACACCAACAGACCTTCTTCCCTGGCGGTTTACGCTTTTCCGTTGCC 823
QY 481 TCTTCAGGAGGATCGTGGCCTGCCACTTCATCAAGCCTCGGACTCTCTTGGACTACTGCC 540
Db 824 GCTTTCAGAAATTTGGGATGCGCATGCCATTTATGAAGCCCGAGGACCTGCTCACTTGGC 883
QY 541 AGGCTCTAGAGAAATCCCGGGGGAGATGTTCCGCTGTCTGA 584
Db 884 AAGCACTGGAGAACTCGAAAGAACAGGACTGCCCTGCTGTCTGA 927

RESULT 6
CEI32434/c
CEI32434/LOCUS
DEFINITION
tiger-gss-dog-17000326239815 Dog Library Canis familiaris genomic,
genomic survey sequence.
ACCESSION
CEI32434
VERSION
CEI32434.1 GI:35233692
KEYWORDS
GSS.
SOURCE
Canis familiaris (dog)
ORGANISM
Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
1 (bases 1 to 766)
Kirkness, B.F., Bafna, V., Halpern, A.L., Levy, S., Remington, K.,
Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and
Venter, J.C.
The dog genome: survey sequencing and comparative analysis
Science 301 (5641), 1898-1903 (2003)
22875432
14512627
Contact: Kirkness EF
The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirknes@tigr.org
Class: shotgun.

```

FEATURES

Location/Qualifiers

1..766
 /organism="Canis familiaris"
 /mol_type="genomic DNA"
 /strain="Standard Poodle"
 /db_xref="taxon:9615"
 /clone_lib="Dog Library"
 /notes="Site 1: BstXI; Libraries were prepared from peripheral blood"

ORIGIN

Query Match 48.4%; Score 282.4; DB 9; Length 766;
 Best Local Similarity 81.2%; Pred. No. 8.7e-74;
 Matches 328; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 1 GGGACATATTCAGAGGATTTCTAGAGCTCTATTACAACTGACCCCTGAAGACCATGA 60
 |||||
 Db 415 GCGACATATTCAGAGGATTTGTGGAGCTTACTTCAATTTGACCCCTGAAGACCATGA 356
 |||||

QY 61 TGGGCTAGAGTGGTCCATCGCTTTTGTCTCAGCGCGGTTTGTGATGAAACAGACT 120
 |||||
 Db 355 TGGGCTATAGAGTGGATCCACCGCTTCTGCTCAGCGGAGCTTTGTGATGAAGCGGACT 296
 |||||

QY 121 CAGACATGTTTCATCATATGTTGACTATCTGACTGAAGTCTTCTGAAGAAAAACAGAA 180
 |||||
 Db 295 CGGACATGTTTGTCAACGCTACTACTGACCGAGCTGCTTCTCAAGAAAAACAGAA 236
 |||||

QY 181 CCAGGTTTTTCACTGGCTTCTTGNAACTCAATGAGTTTCCCATCAGCGAGCCATTTCAGCA 240
 |||||

Db 235 CTCGGTTTTTCAACGGCTTCTTAAACTGACAGTTTCCCATTAGGAGCAAGGCCAAC 176
 |||||

QY 241 AGTGGTTTTCAGTAAATCTGAATATCCGTGGGACAGGTACCCACCATTTCTCTCGGCA 300
 |||||

Db 175 AGTGGTTTTCAGTAAATACGAATACCATGGGATAAGTACCGCGCTTTTGTCTCGGCA 116
 |||||

QY 301 CCGGCTACGTGTTTCTGGGAGCTGGCGAGTCAGGTGTACATGTCTCAAGAGGCTTC 360
 |||||

Db 115 CCGGCTATGTGTTTCCAGTGTATGTGCAAGTCAGGTGTACAACTGCTCTCAGCGCTCC 56
 |||||

QY 361 CATACATTAATCTGAAGACGTGTTTGTGGGCTCTGCGCTCGAA 404
 |||||

Db 55 CGTTCAATTAAGCTGAGACGCTTTGTGGGACTGCGCGGCCA 12
 |||||

RESULT 7

CR564113
 LOCUS 629 bp mRNA linear EST 19-JUL-2004
 DEFINITION XGC-tailbud-head Xenopus tropicalis cDNA clone ThdA01ln08
 5', mRNA sequence.

CR564113.1 GI:50394190

CR564113

EST

Xenopus tropicalis (western clawed frog)

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 Xenopodinae; Xenopus; Silurana.

REFERENCE

1 (bases 1 to 629)
 Croning,M.D.R., Ashurst,J.L., Taylor,R., Garrett,N. and Rogers,J.
 Sanger Xenopus tropicalis EST project 2001 (2004)
 Unpublished (2004)
 Contact: Croning MDR

AUTHORS

Sanger Institute

Hinxton, Cambridgeshire, CB10 1SA, UK

Email: tropesanger.ac.uk

Sanger Xenopus tropicalis EST project 2001

TROPICALIS_SEQUENCE ID: ThdA01ln08.plkSP6

This sequence is from a Xenopus Gene Collection (XGC) library

constructed by Nigel Garrett.

Seq primer: SP6.

FEATURES

Location/Qualifiers

1..629

/organism="Xenopus tropicalis"

source

ORIGIN

Query Match 31.5%; Score 184.2; DB 7; Length 629;
 Best Local Similarity 65.1%; Pred. No. 4e-44;
 Matches 287; Conservative 0; Mismatches 153; Indels 1; Gaps 1;

QY 102 TTTGTGATCAAAACAGACTCAGACATG-TTCATCATGTTGACTATCTGACTGAACGTCT 160
 |||||
 Db 1 TTTGTGATCAAAACAGACTCAGACATGTTTTCGTTAAACACCTTCTACCTGGTCCAACGTCT 60
 |||||

QY 161 TCTGAAGAAAAACAGAAACCAAGGTTTTTCACTGGCTTCTTGAAACTCAATCAGTTTCC 220
 |||||
 Db 61 GGCAGAAAAACCAAGTCTTCTAATTTTTTACTGGTTTCTCAAACTGAACGAGTACCC 120
 |||||

QY 221 CATCAGGAGCCATTCAGCAAGTGGTTTCTCAGTAAATCTGAATATCCGTGGGACAGGTA 280
 |||||

Db 121 GATAAGGAATATCTTCAGCAAGTGGTACGCCAGTAAAGGGAATACCCAGGGGCCAAGTA 180
 |||||

QY 281 CCCACCATCTCTCGGACCGGCTACGTGTTTTCTGGGACGTGGCGAGTCGAGTCAGGTGA 340
 |||||

Db 181 CCTCCATTTTCTCGGAGACTGGATACGTCTTTCTGTAGACGTGCGCCAAAAAGATCCA 240
 |||||

QY 341 CAATGTCTCAAGAGCGTCCCATACATTAACCTGGAAGACGTTGTTGTGGGGCTCTGCGCT 400
 |||||

Db 241 CAACATCTCCAGCAGTCGCGTTTTTCAACTGGAGGACGTCTATTGGGGGTATGTCT 300
 |||||

QY 401 CGAAGGCTGAACATCAGATTCGAGGAGCTCCACTCCAGCCGACCTTTTTTCCAGGGGG 460
 |||||

Db 301 TGACATATTGGACATTCACCTGGAGGAACTTCATACAGAGCAGACATTTCTTTCAGAGAG 360
 |||||

QY 461 CTTACGCTTCTCCGTATGCTCTTCAGGAGGATCGTGGCTGCCACTTCATCAGGCTCG 520
 |||||

Db 361 GCAGTCATTTCTCCGTTTGCAATACAGTAACTTGTACGTCCTCCATGGAGTCAACCAT 420
 |||||

QY 521 GACTCTCTGGACTACTGGCA 541
 |||||

Db 421 TGAGAACATTTGATATCTGGAA 441
 |||||

RESULT 8

AJ507051

LOCUS

DEFINITION

AJ507051

182, mRNA sequence.

ACCESSION

AJ507051

VERSION

AJ507051.1

GI:23034136

KEYWORDS

EST

SOURCE

Mus musculus

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 493)

AUTHORS

Gitton,Y., Dahmane,N., Baik,S., Ruiz i Altaba,A., Neidhardt,L.,

Scholz,M., Hermann,B.G., Kahlem,P., Ben Kahla,A., Schrinner,S.,

Yildirimman,R., Herwig,R., Lehrach,H. and Yaspo,M.L.

A gene expression map of human chromosome 21 orthologs in the mouse

Nature (2002) In press

CONTACT: Kahlem P

Vertebrate Genomics

Max-Planck Institute for Molec. Genet.

Innestrass 73, D-14195 Berlin, GERMANY.

FEATURES

Location/Qualifiers

1..493

source

/mol_type="mRNA"

/db_xref="taxon:8364"

/clone="THdA01ln08"

/dev_stage="tailbud head (stage 28-30)"

/lab_host="Escherichia coli DH10B."

/clone_lib="XGC-tailbud-head"

/note="Vector: pCSI07; Site 1: EcoRI; Site 2: NotI; cDNA was oligo dT primed from sug of poly A+ RNA from tailbud head. EcoRI-NotI cut cDNA was then ligated into pCSI07 with EcoRI at the 5' end and NotI at the 3' end."


```
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="BALB/c"
/db_xref="taxon:10090"
/clone="182"
/clone_lib="Mus musculus BALB/c (Kahlem P)"

ORIGIN

Query Match      29.6%; Score 173; DB 1; Length 493;
Best Local Similarity 73.9%; Pred. No. 9e-41;
Matches 232; Conservative 0; Mismatches 81; Indels 1; Gaps 1;

QY 244 GGTGTGTCAGTAATCTGAATATCCGTCGGACAGGTA-CCACCAATTCCTGTCGGACACC 302
      |||||
Db 52  GGTGTGTCAGTAAGTTCGAATATCCGTCGGACAGGTACCCACCTTTTGTCTGTGGTACT 111
      |||||

QY 308 GCCTACGCTGTTTCTGCGGACGTCGTCAGGTGAGTGTACAAATGTCCTCAAGAGCGTCCCA 362
      |||||
Db 112 GGTATATGCTTTTCCAGCGACGTGCGATCCAAGTATACACGCTCTCAGAGAGCGTTCOG 171
      |||||

QY 368 TACATTAACATGGAAGACGCTGTTGTGGGCTCTGCTCGAAGGCTGAACATCAGATTG 422
      |||||
Db 172 TTCTCAAGCTGGAGGATGTTGTTGTGGGCTCCGCTTGGCCAGACTTAAGATCCGGCGC 231
      |||||

QY 423 GAGGAGCTCCATCCCGACCGACCTTTTTCAGGGGCTTACGCTTCTCGGTATGCTCTC 482
      |||||
Db 232 GAGGAGCTGCACACCAACAGACCTTCTTCCCTGGCGGTTTACGCTTCTCGGTGTCGCGC 291
      |||||

QY 483 TTCAAGGAGATCGTGGCTGCGCTTCATCAAGCTTCGAGCTCTCTTGAGACTTACTGCGCAG 542
      |||||
Db 292 TTTCCANAAATGTGGCATGCCATTTTATGAAGCCCGACGACCTCTCATTACTTGCCAA 351
      |||||

QY 343 GCTCTAGAGAAATC 556
      |||||
Db 352 GCCTGGAGAACTC 365
      |||||

RESULT 9
BP761359 mouse (C57BL/6) pancreatic islet library with
LOCUS recombination-based method Mus musculus cDNA clone mic03025 5',
DEFINITION mRNA sequence.
ACCESSION BP761359.1 GI:50220057
VERSION BP761359
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 492)
JOURNAL Nishimura,M., Yokoi,N., Miki,T., Horikawa,Y., Yoshioka,H.,
CONSTRUCTION Takeda,J., Ohara,O. and Seino,S.
COMMENT Construction of a multi-functional cDNA library specific for mouse
pancreatic islets and its application to microarray
Unpublished (2004)
Contact: Susumu Seino
Division of Cellular and Molecular Medicine
Kobe University Graduate School of Medicine
7-5-1 Kusunoki-cho, Chuo-ku, Kobe, Hyogo 650-0017, Japan
Tel: 81-78-382-5360
Fax: 81-78-382-5370
Email: seino@med.kobe-u.ac.jp.
Location/Qualifiers
1. .492
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="BALB/c"
/db_xref="taxon:10090"
/clone="mic03025"
/sex="male"
/tissue_type="pancreatic islet"
/dev_stage="adult"

FEATURES
source
1. .492
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="BALB/c"
/db_xref="taxon:10090"
/clone="mic03025"
/sex="male"
/tissue_type="pancreatic islet"
/dev_stage="adult"

/clone_lib="mouse (C57BL/6) pancreatic islet library with
recombination-based method"

ORIGIN

Query Match      26.3%; Score 153.8; DB 5; Length 492;
Best Local Similarity 79.5%; Pred. No. 5.8e-35;
Matches 182; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 1 GGGACATATCCAGAAGATTTCCTAGACGCTCTATTACATCTGACCTGAGACCATGA 60
      |||||
Db 251 GCGACATATCCAGAAGGACTTCAAGGATGCTCTACTTTCAACCTGACCTGAAAACCATGA 310
      |||||

QY 61 TGGGCATAGATGGTCCATCGCTTTTGTCTCAGGGGCGCTTTGTGATGAAACAGACT 120
      |||||
Db 311 TGGGTATGGAATGGTCTACCACCTTTTGTCTCAGACGCTTACGTGATGAAACGGACT 370
      |||||

QY 121 CAGACATGTTTCATCAATGTTGACTATCTGACTGAACTGCTTCTGAAGAAAAACAGAA 180
      |||||
Db 371 CTGACATGTTTGTGAATGTTGGCTTCTGACGGAACCTGCTCTTAAGAAAAACCAACGA 430
      |||||

QY 181 CCAGGTTTTCACCTGGCTTCTTGAACTCATGAGTTTCCCATCAGGCA 229
      |||||
Db 431 CCAGGTTTTCACAGGCTACATTAAGCCCCACGACTTTTCCATTGGCA 479
      |||||

RESULT 10
BP302370 tcay Oncorhynchus mykiss cDNA clone tcay0008b.1.08 Sprim,
LOCUS mRNA sequence.
DEFINITION BP302370.2 GI:42614467
ACCESSION BP302370
VERSION BP302370
KEYWORDS Oncorhynchus mykiss (rainbow trout)
SOURCE Oncorhynchus mykiss
ORGANISM Oncorhynchus mykiss
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Actinopterygii; Neopterygii; Teleostei; Euteleostei;
TITLE Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
JOURNAL 1 (bases 1 to 626)
COMMENT Govoroun,M., Guiguen,Y. and Le Gac,F.
Construction and primary characterization of normalized cDNA
libraries in rainbow trout, Oncorhynchus mykiss
Unpublished (2003)
On Apr 7, 2003 this sequence version replaced gi:29583015.
Contact: Guiguen Y
INRA - SCRIBE
Campus de beaulieu, RENNES cedex, 35042, France
Tel: 02.23.48.50.09
Fax: 02.23.48.50.20
Email: Yann.Guiguen@beaulieu.rennes.inra.fr
Sequence cleaned of vector, adaptor and repetitions. Contact us
at signasupport@jouy.inra.fr to obtain the chromatogram of this
sequence.
Plate: 0008 row: 1 column: 8
Seq primer: M13R
Location/Qualifiers
1. .626
/organism="Oncorhynchus mykiss"
/mol_type="mRNA"
/db_xref="taxon:8022"
/clone="tcay0008b.1.08"
/tissue_type="adipose tissue, blood, brain,
differentiating gonads, gills, interrenal, intestine,
kidney, liver, muscle, ovary, pituitary, testis"
/dev_stage="from embryos to adults"
/lab_host="DH10B"
/clone_lib="tcay"
/clone="Vector: pT73D-pac; Rainbow trout multi-tissues -
normalized + 1 subtraction (tcay) ; Clone distribution :
AGENAE Resource centre. Francois PIUMI,
Francois.Piumi@jouy.inra.fr, INRA, CEA Radiobiologie et
Etude du genome (LREG), Domaine de Vilvert, 78352,
Jouy-en-Josas cedex, France, +33 (0) 1.34.65.28.02, +33
```


ORIGIN		FEATURES		Location/Qualifiers	
				1. .901	
				/organism="Oryzias latipes"	
				/mol_type="mRNA"	
				/strains="Hd-rR"	
				/db_xref="taxon:8090"	
				/sex="mixture of female and male"	
				/tissue_type="whole embryo"	
				/dev_stage="organogenesis stage 35"	
				/clone_lib="MP015DA cdna"	
ORIGIN					
				Query Match 22.4%; Score 131; DB 4; Length 901;	
				Best Local Similarity 53.7%; Pred. No. 5.4e-28;	
				Matches 297; Conservative 0; Mismatches 250; Indels 6; Gaps 1;	
QY		1		GGGACATATCCAGAGGATTTCTTAGAGCTTATTAACAATCTGACCTGAGGACCATGA	
DB		675		GTGACATCATCCACAGGACTTTTGGATTCCTATAAAAAATCTTACTATANAAGACTGA	
QY		61		TGGGATAGATGGGTCCATCGCTTTTGTCTCAGCGCGCTTTGTGATGAAACAGACT	
DB		615		TGGGATAGATGGGTGGGAGTGCATGCCAAGCGGCAAGTTACGTCATGAAACAGACA	
QY		121		CAGACATGTTTCAATCAATGTTGACTATCTGACTGAACTGCTTCTGAAGAAA-----AACA	
DB		555		GTGACATGTTTCTGCAACAGAGTACCTTATTTATAAGCTGCTGAAGCCAGAACTGAAG	
QY		175		GAAACAAACGAGTTTTCATGCTGCTTCTGAAATCTGAATGAGTTTCCATCAGGACCCAT	
DB		495		CCAAACCAACTACTTACAGGAAACAACATGAGAGGCTTTGCACCAACCGGAATAAAA	
QY		235		TCAGCAAGTGGTTTCTGAGTAATCTGAATATCTGTTGGGACAGGTACCCACCATCTGCT	
DB		435		ACAGCAAGTGGTCAATGTTCCCGGAGCTGTACCCGAGGAAAGTACCCCATTTTTCCT	
QY		295		CCGGCACCGGCTACGTTGTTTCTGCGAGCTGCGAGTCCAGGTGAGTACAAATGTTCTCAAGA	
DB		375		CTGGAATCGGCTATGTTTCTCGGAGACATGCGCGGAAATCTTACAGACTTCTCTTGA	
QY		355		GGTCCCATACATTAACACTGGAAGAGCTGTTTGTGGGCTCTGCTCGAAAGGCTGAACA	
DB		315		GCACCCCGGCTGACCTGACCTGGAGGAGTATAGCTGGGAATATGCTTGGCCAGACTCAGAA	
QY		415		TCAGATTGGAGGAGCTCCACTCCAGCGGAGCTTTTTCAGGGGGCTTACGCTTCTCCG	
DB		255		TCGAGCCCAACCGCGCGGCTTCTGTTCAACCACTGGGCGGCTCTCTTACTCCA	
QY		475		TATGCTCTTTCAGGAGGATCGTGGGCTGCGACTTTCATCAAGGCTCGGACTCTCTTGGACT	
DB		195		GCTGCAAGTACAGCCACCTGATAACATCACATGGGTTCAGTCCCAATGAATACTGAAT	
QY		535		ACTGGAGGCTCT 547	
DB		135		ACTGGCATCATCT 123	
RESULT 12					
AY401366				981 bp DNA linear GSS 15-DEC-2003	
LOCUS				Mus musculus B3GALT1 gene, VIRTUAL TRANSCRIPT, partial sequence,	
DEFINITION				genomic survey sequence.	
ACCESSION				AY401366	
VERSION				AY401366.1 GI:39757355	
KEYWORDS				GSS.	
SOURCE				Mus musculus (house mouse)	
ORGANISM				Mus musculus	
REFERENCE				Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS				Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
				1 (bases 1 to 981)	
				Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,	
				Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,	

(O) 1.34.65.22.73"		Query Match 25.4%; Score 148.2; DB 5; Length 626;	
		Best Local Similarity 56.0%; Pred. No. 3.1e-33;	
		Matches 304; Conservative 0; Mismatches 233; Indels 6; Gaps 1;	
QY		3	
DB		83	
QY		63	
DB		143	
QY		123	
DB		203	
QY		183	
DB		263	
QY		237	
DB		323	
QY		297	
DB		383	
QY		357	
DB		443	
QY		417	
DB		503	
QY		477	
DB		563	
QY		537	
DB		623	
RESULT 11			
BJ740020/c		901 bp mRNA linear EST 10-MAR-2004	
LOCUS		BJ740020 MP015DA cdna Oryzias latipes cdna clone MP015DA012002 3',	
DEFINITION		mRNA sequence.	
ACCESSION		BJ740020	
VERSION		BJ740020.1 GI:45337290	
KEYWORDS		EST.	
SOURCE		Oryzias latipes (Japanese medaka)	
ORGANISM		Oryzias latipes	
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
		Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;	
		Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;	
		Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.	
REFERENCE		1 (bases 1 to 901)	
AUTHORS		Kohara, Y., Shin-i, T., Kimura, T., Narita, T., Jindo, T. and Takeda, H.	
TITLE		Medaka EST Project in Takeda's lab	
JOURNAL		Unpublished (2001)	
COMMENT		Contact: Tadao Shin-i	
		Center For Genetic Resource Information	
		National Institute of Genetics	
		1111 Yata, Mishima, Shizuoka 411-8540, Japan	
		Tel: 81-559-81-6856	
		Fax: 81-559-81-6855	
		Email: tshini@genes.nig.ac.jp.	

misc_feature
/clone="9530002C17"
/sex="male"
/issue_type="urinary bladder"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
284_1263
/note="UDP-Gal:betaGlcNAc beta 1,3-galactosyltransferase,
polypeptide 1 (MGDI)GI:1349403, GB|NM_020283, evidence:
BLASTN, 99%, match=980)
putative"
polya_signal
2682..2687
/note="putative"
polya_site
2705
/note="putative"

ORIGIN

Query Match 21.0%; Score 122.8; DB 3; Length 2705;
Best Local Similarity 59.3%; Pred. No. 2.3e-25;
Matches 249; Conservative 0; Mismatches 162; Indels 9; Gaps 2;
QY 3 GACATTATCCAGAGGATTTCTAGACGCTATTACAACTCTGACCCCTGAAGACCATGATG 62
DB 697 GACATCATCTGGAGGACTTCATTGACTCTACCAACAATCTCACCTCTCAAACCTTAATG 756
QY 63 GGCATAGATGGGTCCATCGCTTTTCTCTCAGCGCGCGTTTGTGTGATGAAGAAACAGACTCA 122
DB 757 GGGATGAGATGGGTGGCCACTTCTGTCTCAAAAGCCAAGTAGCTCATGAAGAAACCGACAGT 816
QY 123 GACATGTTTCATCAATGTTGACTATCTGACTGAAGTCTCTTGAA-----GAAACACAGA 176
DB 817 GACATTTTGTGAACATGACAACTTATTATAAACTCTGAAACCCCTCTACCAAGCCA 876
QY 177 ACAACGAGGTTTTTCACTGGCTTCTTGAACACTCAATGAGTTTCCCATCAGGAGGCCATTC 236
DB 877 AGNAGAGGATTTTCACTGGTAGCTCA---TCNACGGGGGCCAATCAGGATGTCCGC 933
QY 237 AGCAAGTGGTTGTCTAGTAAATCTGAAATATCGGTGGGACAGGTACCACCAATCTGTCTCC 296
DB 934 AGTAAGTGTATATGCTAGAGATTTGTACCCCTGCAGCAAACTACCCACCGTTCTGTTC 993
QY 297 GGCACGGCTAGCTGTTTCTGGCAGGTGGCAGTCAAGTGTACAAATGTCTCCAAGAGC 356
DB 994 GGGAGTGGCTATATCTTTTCCGCTGATGTGGCTGAACACTCAITTTACAAGACCTCGCTCC 1053
QY 357 GTTCCCATACATAAACTGGAAGACGTTGTTGTGGGCTCTGCTCGAAAGGCTGAACATC 416
DB 1054 ACCAGGCTGTTCAATCTTGAAGATGTGTACGTGGGACTGTGTTCTCGAAAGCTTGGCATC 1113

RESULT 14
BU133295
LOCUS 603122104F1 CSEQCHL22 Gallus gallus cDNA clone Chest85h2 5', mRNA
DEFINITION sequence.
ACCESSION BU133295
VERSION BU133295.1 GI:25345664
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus

REFERENCE 1 (bases 1 to 685)
AUTHORS Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
TITLE A Comprehensive Collection of Chicken cDNAs
JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)
MEDLINE 22335534
PUBMED 12445392
COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)

PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.
Location/Qualifiers

FEATURES
source

1..685
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hixex"
/db_xref="taxon:9031"
/clone="CHEST85h2"
/dev_stage="36"
/lab_host="DH10B"
/clone_lib="CSEQCHL22"
/note="Organ: heads; Vector: pBluescript II KS(+); Site 1:
EcoRI; Site 2: NotI; Modification of pBluescript II KS(+)
[Stratagene] vector to accommodate cDNA produced with the
T-trimmed protocol (Construction of uni-directionally
cloned cDNA libraries from messenger RNA for improved 3'
end DNA sequencing by Glenn Fu, et al. U.S. Patent
6,387,624). Cut pBluescript II KS(+) with NotI and EcoRI.
Ligate in double stranded adaptor containing BglI and
BamHI sites [5'ggcggcgtagcccgccggaacgaagaaag]
[5'aattcttttttcggatccggggctgcacgc]"

ORIGIN

Query Match 20.8%; Score 121.2; DB 5; Length 685;
Best Local Similarity 59.0%; Pred. No. 4.6e-25;
Matches 248; Conservative 0; Mismatches 163; Indels 9; Gaps 2;
QY 3 GACATTATCCAGAGGATTTCTAGACGCTATTACAACTCTGACCCCTGAAGACCATGATG 62
DB 123 GACATTATTTAGAGGATTTTATTGACTCTCTATCACTCTGAAACATTAAATG 182
QY 63 GGCATAGATGGGTCCATCGCTTTTGTCTCAGCGCGCGTTTGTGTGATGAAGAAACAGACTCA 122
DB 183 GGGATGAGTGGGTGAGCAACATATTGTTCAAAAGCGAAGTAGCTTATGAAGACAGACAGT 242
QY 123 GACATGTTTCATCAATGTTGACTATCTGACTGAACCTGTTCTCAAGAAACAA-----GA 176
DB 243 GATATTTTAGTAATATATGATTAATCTTATTATAGTCTCTCAACCTTAACCAAGCCA 302
QY 177 ACAACGAGGTTTTTCACTGGCTTCTTGAAACTCAATGAGTTTCCCATCAGGAGCCATTC 236
DB 303 AGGAGAGGCTACTTCACTGGTTACGTTA---TCAACGGAGGGCCAAATACGAGACGTTCCG 359
QY 237 AGCAAGTGGTTTGTCTAGTAAATCTGAAATATCGTGGGACAGGTACCACCAATCTCTCTCC 296
DB 360 AGCAAGTGGTACATGCCAGAGATTTGTATCTCTGACAGCAATTTACCCACCTTCTGTTCC 419
QY 297 GGCACGGCTAGCTGTTTCTCGCAGCGTGGCAGTCAAGTGTACAAATGTCTCCAAGAGC 356
DB 420 GGCATGGCTACATTTTTCAGCAGATGTAGCAGAACTGATTTACAAACCTCCCTTCCAC 479
QY 357 GTTCCCATACATTAAACTGGAAGACGTTGTTGTGGGCTCTGCTCGAAAGGCTGAACATC 416
DB 480 ACCAGACTCTCCATCTCGAAGACGTTGATGTGCGACTCTGTCTTCGGAAGCTGGGCATC 539

RESULT 15
BU350312
LOCUS 603527845F1 CSEQCHN69 Gallus gallus cDNA clone Chest47712 5', mRNA
DEFINITION sequence.
ACCESSION BU350312
VERSION BU350312.1 GI:25858313
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 842)

AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
Pong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)
22335534
12445392
Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.

FEATURES
source

1..842
Location/Qualifiers
/organism="Gallus gallus"
/mol_type="mRNA"
/strains="Compton Line 151"
/db_xref="taxon:9031"
/clone="CHST477112"
/sex="Female"
/tissue_type="cerebellum"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="CSQCHN69"
/note="Organ: brain; Vector: pBluescript II KS(+); Site 1:
EcoRI; Site 2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunted, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."

ORIGIN

Query Match 20.8%; Score 121.2; DB 5; Length 842;
Best Local Similarity 59.0%; Pred. No. 4.9e-25;
Matches 248; Conservative 0; Mismatches 163; Indels 9; Gaps 2;
3 GACATTATCCAGAAGGATTTCTAGACGTCATTACAAATCTGACCCCTGAAGACCATGATG 62
21 GACATTATTTGTAGAGGATTTTATTTGACTCTCTATCATACCTCCTCTGAAAACATTATG 80
63 GGCATAGATGGGTCCATCGCTTTTGTCTCAGGGCGGCTTTGTGATGAAACACACTCA 122
81 GGGATGAGGTGGGTAGCAACATATGTTTCAAAGCGAAGTAGCTTTATGAAGACAGACAGT 140
123 GACATGTTTCATCAATGTTTCACTACTGCTGACTGCTTCTGAAGAAACA- ----GA 176
141 GATAATTTTGTAAATATGATATCTTATTTAAGTCTCTCAAACCTAACACCAAGCCA 200
177 ACAACCAAGGTTTTTCACTGGCTTCTTGAACACTCAATGAGTTTCCCATCAGGCAGCCATT 236
201 AGGAGAAGGTACTTCACTGTTACGTTA- - -TCAACGGAGGGCCAAATACGAGCGTTGCG 257
237 AGCAAGTGGTTTGTCTAGTAAATCTGAAATATCCGTGGGACAGGTACCCACCATTTCTGCTCC 296
258 AGCAAGTGGTACATGCCAGAGATTTGTATCTCTGACAGCAATTTACCACCCCTTCTGTTCCG 317
297 GGCACCGGTACGTGTTTCTCGCAGCTGGCGAGTCAAGTGTACAAATGCTCTCCAGAGC 356
318 GGCACGTGGCTTACATTTTTTTCAGCAGATGAGCAGAACTGATTTACAAAACCTCCCTTCA 377
357 GTCCCATACATTAACCTGGAAGACGTGTTTGTGGGGCTCTGGCTCGAAGAGGCTGAACATC 416
378 ACCAGACTCCTCCATCTCGAAGACGTGTATGTCGACTCTGTCTTCGGAAGCTGGGCATC 437

Search completed: September 22, 2005, 16:56:10
Job time : 2583.06 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2005, 12:15:17 ; Search time 121.581 Seconds
(without alignments)
7859.644 Million cell updates/sec

Title: US-10-777-828-8_COPY_428_1011

Perfect score: 584

Sequence: 1 gggacattatccagaagat.....aagattgtccgctgtctga 584

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PTCUS_COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	584	100.0	1011	US-09-831-630-8	Sequence 8, Appli
2	80.8	13.8	1134	US-09-482-180A-3	Sequence 3, Appli
3	56.2	9.6	1191	US-09-459-133-3	Sequence 3, Appli
4	53	9.1	1167	US-09-459-133-14	Sequence 14, Appli
5	52.4	9.0	1464	US-09-774-528-430	Sequence 430, Appli
6	48	8.2	1420	US-09-482-180A-1	Sequence 1, Appli
7	48	8.2	22374	US-09-949-016-14938	Sequence 14938, A
8	46	7.9	125902	US-09-949-016-13715	Sequence 13715, A
9	40.2	6.9	1434	US-09-055-097-2	Sequence 2, Appli
10	40.2	6.9	1434	US-09-373-902-2	Sequence 2, Appli
11	40.2	6.9	1640	US-09-949-016-5696	Sequence 5696, Ap
12	40.2	6.9	5640	US-09-949-016-17438	Sequence 17438, A
13	40.2	6.9	8436	US-09-949-016-13963	Sequence 13963, A
14	39.2	6.7	748	US-09-270-767-444	Sequence 444, App
15	39.2	6.7	748	US-09-270-767-15726	Sequence 15726, A
16	38.6	6.6	17731	US-09-949-016-16365	Sequence 16365, App
17	37	6.3	2359	US-09-620-312D-699	Sequence 699, App
18	37	6.3	2426	US-09-620-312D-697	Sequence 697, App
19	36.6	6.3	1098	US-09-716-793A-3	Sequence 3, Appli
20	35.8	6.1	374	US-09-270-767-1385	Sequence 1385, Ap
21	35.8	6.1	374	US-09-270-767-16667	Sequence 16667, A
22	35.4	6.1	1279	US-09-902-540-6180	Sequence 6180, Ap
23	35.4	6.1	1959	US-09-902-540-313	Sequence 313, App
24	35	6.0	2557	US-08-464-954A-1	Sequence 1, Appli
25	34.4	5.9	861	US-09-347-613C-3	Sequence 3, Appli
26	34.4	5.9	861	US-09-662-183A-3	Sequence 3, Appli
27	34.2	5.9	1100	US-09-248-335-53	Sequence 53, Appli

ALIGNMENTS

RESULT 1

US-09-831-630-8
; Sequence 8, Application US/09831630
; Patent No. 6800468
; GENERAL INFORMATION:
; APPLICANT: Clausen, Henrik
; APPLICANT: Amado, Margarita
; TITLE OF INVENTION: UDP-GALACTOSE: BETA-N-ACETYL-GLUCOSAMINE BETA 1,3
; FILE REFERENCE: 7188-157
; CURRENT APPLICATION NUMBER: US/09/831,630
; CURRENT FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 1011
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (79)..(1008)
US-09-831-630-8

Query Match 100.0%; Score 584; DB 4; Length 1011;
Best Local Similarity 100.0%; Pred. No. 1.2e-183;
Matches 584; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGGACATTATCCAGAAAGATTTCCTAGACGCTATTACAACTGACCCCTGAAGACCATGA 60
Db 428 GGGACATTATCCAGAAAGATTTCCTAGACGCTATTACAACTGACCCCTGAAGACCATGA 487
Qy 61 TGGGCATAGAAATGGTCCATCGCTTTTCTCCTCAGCGCGGTTTGTGTGATGAAAAAGAGCT 120
Db 488 TGGGCATAGAAATGGTCCATCGCTTTTCTCCTCAGCGCGGTTTGTGTGATGAAAAAGAGCT 547
Qy 121 CAGACATGTTTCATCAATGTTGACTATCTGACTGAACCTGCTTCTGAAGAAAAACAGACAA 180
Db 548 CAGACATGTTTCATCAATGTTGACTATCTGACTGAACCTGCTTCTGAAGAAAAACAGACAA 607
Qy 181 CCAGGTTTTTCACCTGGCTTCTTGAACATCAATAGTTCCTCCTCAGGAGGAGCCATTTCAGCA 240
Db 608 CCAGGTTTTTCACCTGGCTTCTTGAACATCAATAGTTCCTCCTCAGGAGGAGCCATTTCAGCA 667
Qy 241 AGTGTGTTTGTGAGTAAATCTGAATATCCGTGGGACAGTACCCACATTTGTCTCCGGCA 300
Db 668 AGTGTGTTTGTGAGTAAATCTGAATATCCGTGGGACAGTACCCACATTTGTCTCCGGCA 727
Qy 301 CCGGTAGCTGTTTTCGGCGACGCTGGCGAGTCAGGTGTACATGTTCTCCAGAGCGTCC 360
Db 728 CCGGTAGCTGTTTTCGGCGACGCTGGCGAGTCAGGTGTACATGTTCTCCAGAGCGTCC 787

```
QY 361 CATACATTAAACTGAAGACGTTGTTGGGGCTCTGCTCGAAAGGCTGAACATCAGAT 420
Db 788 CATACATTAAACTGAAGACGTTGTTGGGGCTCTGCTCGAAAGGCTGAACATCAGAT 847
QY 421 TGGAGGAGCTCCACTCCAGCGGACCTTTTTCAGGGGGCTTACGGCTTCTCCGATATGCC 480
Db 848 TGGAGGAGCTCCACTCCAGCGGACCTTTTTCAGGGGGCTTACGGCTTCTCCGATATGCC 907
QY 481 TCTTCAGAGAGATCGTGCGCTGCCACTTTCATCAAGCCTCGGACTCTCTTGGACTACTGGC 540
Db 908 TCTTCAGAGAGATCGTGCGCTGCCACTTTCATCAAGCCTCGGACTCTCTTGGACTACTGGC 967
QY 541 AGGCTCTAGAGAAATCCCGGGGGAGAGATTCTCGCGCTGTCTGA 584
Db 968 AGGCTCTAGAGAAATCCCGGGGGAGAGATTCTCGCGCTGTCTGA 1011

RESULT 2
US-09-482-180A-3
; Sequence 3, Application US/09482180A
; Patent No. 6361985
; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell C.
; APPLICANT: Yamamoto, Gayle
; APPLICANT: Gao, Zeren
; APPLICANT: Whitmore, Theodore E.
; APPLICANT: Jaepers, Stephen
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING
; FILE REFERENCE: BETA-1.3-GALACTOSYLTRANSFERASE HOMOLOG, ZN5SP6
; CURRENT APPLICATION NUMBER: US/09/482,180A
; CURRENT FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/115,721
; PRIOR FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 1134
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: degenerate sequence
; NAME/KEY: Variation
; LOCATION: (1)...(1134)
; OTHER INFORMATION: n is any nucleotide
; NAME/KEY: misc feature
; LOCATION: (1)...(1134)
; OTHER INFORMATION: n = A,T,C or G
US-09-482-180A-3

Query Match 13.8%; Score 80.8; DB 3; Length 1134;
Best Local Similarity 32.2%; Pred. No. 3.8e-16;
Matches 134; Conservative 61; Mismatches 221; Indels 0; Gaps 0;

QY 3 GACATATCCAGAGGATTTCTAGACGCTCTATACATCTGACCTCGAAGACCATGATG 62
Db 598 GAYATHYNTCAATGGGATTTACNGARGAYTTTAAAYTNACNYTNAARGARYTNCA 597
QY 63 GGCATAGAAATGGTCCATCGCTTTTGTCTCAGGGGGGTTGTGATGAAACACAGATCA 122
Db 598 YTNCAAMGNTGGTNGTNGCNGTGYCNCAAGNCAYTTATGTYTNAARGNGAYGAY 657
QY 143 GACATGTTTCATCAATGTTGACTATCTGACTGAACTGCTTCTTGAAGAAAAACAGAACACC 182
Db 618 GAYGNTTGTGNCAYGTGNCNAAAYGTNTGARTTYTNGAYGNGTGGGAYCCNGCNCAR 717
QY 143 AGGTTTTTCACTGGCTTCTTGAACATCAATGATTTTCCCATCAGCAGCCATTAGCAGAG 242
Db 718 GAYYNTYNTGNGNGAYGTNATHMGNCAAGCNYTNCNAAAYMGNAAYACNNAARGTNAAR 777
QY 243 TGGTTTGTGAGTAAATCTGAATATCCGTGGGACAGGTACCCACCATCTCTGCTCCGGCACC 302
```

```
Db 778 TAYTVAATHCCNCNWSNATGTATMGNMGNACNCAYTAYCCNCCNTAYCGNGNGNGNGN 837
QY 303 GGCTACGCTGTTTCTGCGGAGTGGCGAGTCAAGGTGTACAATGTCTCCAAGAGCGTCCCA 362
Db 838 GENTAYGTNATGNSMNGMNCNACNGTMMGNMGNNTYTCARGCNATHATGGARGAYGCGNGAR 897
QY 363 TACATTAAACTGAAGACGTTGTTGGGGCTCTGCGCTCGAAAGGCTGAACATCAG 418
Db 898 YTNVWNSNATHGAYGAYGNTTYGTGNGNATGTGYTNNMGNMGNNTYNGNGNYTNWS 953

RESULT 3
US-09-459-133-3
; Sequence 3, Application US/09459133
; Patent No. 8416988
; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell C.
; APPLICANT: Yamamoto, Gayle
; APPLICANT: Jaepers, Stephen R.
; APPLICANT: Gao, Zeren
; TITLE OF INVENTION: BETA-1,3-GALACTOSYLTRANSFERASE HOMOLOGS
; FILE REFERENCE: 98-77
; CURRENT APPLICATION NUMBER: US/09/459,133
; CURRENT FILING DATE: 1999-12-10
; PRIOR APPLICATION NUMBER: 60/111,697
; PRIOR FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 1191
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Degenerate sequence
; NAME/KEY: misc feature
; LOCATION: (1)...(1191)
; OTHER INFORMATION: n = A,T,C or G
US-09-459-133-3

Query Match 9.6%; Score 56.2; DB 3; Length 1191;
Best Local Similarity 28.5%; Pred. No. 6e-08;
Matches 161; Conservative 73; Mismatches 322; Indels 9; Gaps 2;

QY 1 GGGACATATCCAGAGATTTCTAGACGCTCTATTACATCTGACCTGAGACCATGA 60
Db 620 SNGAYYNTYNTYNTGGGAYTTTYYTNGAYTNCNTTYYAAYCARACNYTNAARGAYTNY 679
QY 61 TGGGCATAGAAATGGTCCATCGCTTTTGTCTCAGGGGGGTTGTGATGAAAAACAGACT 120
Db 680 TNYTNYTNGCNTGGYTNNGMNGNCAYTYCCNACNGTNWSNTTYGTNYTNGMNGCNCARG 739
QY 121 CAGACATGTTTCATCAATGTTGACTATCTGACTGAATCTGCT-----TCTGAAGAAAAACA 174
Db 740 AYGAYGCVNTTGTNCAVACNCCNGCNYTNTYTNMNGNCAYTYCCNACNGTNYTNCNCCNGCW 799
QY 175 GAACACACAGGTTTTCACCTGGCTTCTTGAAACTCAATGAGTTTCCCATCAGGCGCAT 234
Db 800 SNGCMNGWSNYTNTYATYTNNGNGARGTNTTYACNCAAGCNCATGCCNYTNNMNAARCCNG 859
QY 235 TCACAAAGTGTGTTGTCAGTAAATCTGAATATCCGTGGGACAGGTACCCACATCTGCT 294
Db 860 GNGNCCNTTYATGTCNCGARGNSNTTYTYGARGNG---GNTAYCCNCCNTAYGNCW 916
QY 295 CCGGCACCGGCTACGTGTTTCTGGCGACGCTGGCGAGTCAAGGTGTACAATGTCTCCAAGA 354
Db 917 SNGGNGGNGNTAYGTNATHGNGMNGNYTNGCNCCTGGYTNNTMNGMNGCNGCNCW 976
QY 355 CGGTCCCATACATTAATACTGAAGACGCTGTTGTGGGCTCTGCTCGAAAGGCTGAACA 414
Db 977 GNGTNGCNCNTTYCCNTTYGARGAYGNTAYACNGGNYTNTYATMNGMNGCNYTNGNY 1036
QY 415 TCAGATTGGAGGAGCTCCACTCCCGGCGGACCTTTTTTCCAGGGGGCTTACGCTTCTCCG 474
```

Db 1037 TNGTNCNCARCNCAAYCCNGGNTTYYTNACNGCNTGGCCNGCNGAYMGNACNGCNGAYC 1096
Qy 475 TATGCTCTTTCAGGAGGATCGTGGCTGCCTTCAATCAAGCTCGGACTCTCTTGGACT 534
Db 1097 ATYGYCNTTYNGNAAYYTNTYNTGNTMGNCNTYNGGCCNCARCGNWSNATHMGNY 1156
Qy 535 ACTGGCAGCTCTAGAGAAATCCCG 559
Db 1157 TNTGGAARCARYTNCARGAYCCNMG 1181

RESULT 4

US-09-459-133-14
; Sequence 14, Application US/09459133
; Patent No. 641698
; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell C.
; APPLICANT: Yamamoto, Gayle
; APPLICANT: Jasper, Stephen R.
; APPLICANT: Gao, Zeren
; TITLE OF INVENTION: BETA-1,3-GALACTOSYLTRANSFERASE HOMOLOGS
; FILE REFERENCE: 98-77
; CURRENT APPLICATION NUMBER: US/09/459,133
; CURRENT FILING DATE: 1999-12-10
; PRIOR APPLICATION NUMBER: 60/111,697
; PRIOR FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 1167
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: degenerate sequence
; NAME/KEY: misc_feature
; LOCATION: (1)..(1167)
; OTHER INFORMATION: n = A,T,C or G
US-09-459-133-14

Query Match 9.1%; Score 53; DB 3; Length 1167;
Best Local Similarity 32.4%; Pred. No. 6.9e-07;
Matches 132; Conservative 53; Mismatches 213; Indels 9; Gaps 2;
Qy 1 GGGCATTATCAGAAGGATTTCTAGAGCTATTACAACTGACCTCGAAGACCATGA 60
Db 596 GNGAYTNTYNTGGAYTYYTNGAYTNCNTAYAYMGNACNTYNAARGAYTNY 655
Qy 61 TGGGCATAGAATGGTCCATCGCTTTTCTCTCAGCGCGCGTTTGTGATGAAACAGACT 120
Db 656 TNYTNTACNTGGYTNSNCAYCAYTGCNGAYGTWAAYYTYGTYTNCARGTNCARG 715
Qy 121 CAGACATGTTCAATCAATGTTGACTATCTGACTGAATCTGCTCTGAAGAA-----AAACA 174
Db 716 AYGAYGCTTGTNCAYATHCNGCNYTNYNGARCAYYTNCARACNTYNCNCNCAGT 775
Qy 175 GAACAACAGGTTTCTACTGCTTCTTGAACCTCAATGAGTTTCCCATCAGGAGCCAT 234
Db 776 GGGCNGMWSNYTNTAYTNGNGARATHTTYACNARGCNAARCCNTYNMGNARCCNG 835
Qy 235 TCAGCAAGTGGTGTCTAGTAAATCTGAATATCCGTGGACAGGTPACCCACCATCTGCT 294
Db 836 GNGNGCNYTAYTNCNARACNTYYTYGARGNGAY---TAYCNGNTAYGCNW 892
Qy 295 CCGGACCGGCTACGTGTTTCTCGGACGCTGGAGTCAATGTCTTCAAGA 354
Db 893 SNGGNGGNGNTAYGTNATHWSNGMNGNYNGCNCNTGGYTYTNCARGCNGCNGCM 952
Qy 355 GCGTCCCATACATTAACCTGAGACGCTTGTGGGCTCTGCCTC 401
Db 953 GNGTNGCNCNTTYCCNTTYGAYGAYGTNTAYACNGGNTTYGTYY 999

RESULT 5

US-09-774-528-430
; Sequence 430, Application US/09774528
; Patent No. 6743619
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
; APPLICANT: Yang, Yonghong
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wang, Dunrui
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6743619el Nucleic Acids and
; FILE REFERENCE: 802
; CURRENT APPLICATION NUMBER: US/09/774,528
; CURRENT FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 441
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 430
; LENGTH: 1464
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1464)
US-09-774-528-430

Query Match 9.0%; Score 52.4; DB 4; Length 1464;
Best Local Similarity 47.5%; Pred. No. 1.2e-06;
Matches 189; Conservative 0; Mismatches 206; Indels 3; Gaps 1;

Qy 3 GACATTATCCAGAGGATTTCTAGACGCTATTACATCTGACCTGAAGACCATGATG 62
Db 613 GATATAATTGAAGGAATCTTCTTGACAGAGTTCTGAGAACCAACCTGAAGATCATTCGA 672
Qy 63 GGCATAGAAATGGGTCCATCGCTTTTGTCTCAGCGCGGTTTGTGATGAAACAGACTCA 122
Db 673 ATGATACAGTGGGCTGTGGCTTTCTGCCCTAATGCCCCTGTTCTCAAGTGGATGNA 732
Qy 123 GACATGTTCAATCAATGTTGACTATCTGACTGAACCTGCTTCTGAAGAAACAGAACACC 182
Db 733 GAGAGCTTTGTCAATCTACCAAGCTTGGTAGACTATCTTCTCAATCTGAAGAACACCTA 792
Qy 183 AGGTTTTCATCGCTTCTTGAACCTCAATGAGTTT---CCCATCAGGAGCCATTTCAGC 239
Db 793 GAAGATATCTATGTAGGAAGAGTTTCTTTCATCAGGTTACACCCCAATAGAGATCCTCAGAAC 852
Qy 240 AAGTGGTTTGTTCAGTAAATCTGAATATCCGTGGGACAGGTACCCACCATCTGCTCCGCG 299
Db 853 AGAGACTTTGTCCCTCTTAGTGAGTACCAGAAATACTACCCAGATTACTGAGTGT 912
Qy 300 ACCGCTACGTGTTTCTGGCGAGCTGGCGAGTCAAGTGTACAAATGTCTTCCAAGAGCGTC 359
Db 913 GAGGCTTTATAATGTCCCAAGATGTGGCTCGAATGATGTATGTGTTTTCAGGAAGTA 972
Qy 360 CCATACATTAACCTGGAAGACGTGTTTGTGGGCTCTG 397
Db 973 CCCATGATGGTCCAGCTGATGTGTTTGTAGGAATTTG 1010

RESULT 6

US-09-482-180A-1
; Sequence 1, Application US/09482180A
; Patent No. 6361985


```

; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell C.
; APPLICANT: Yamamoto, Gayle
; APPLICANT: Gao, Zeren
; APPLICANT: Whitmore, Stephen
; APPLICANT: Jaspers, Stephen
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING
; TITLE OF INVENTION: BETA-1,3-GALACTOSYLTRANSFERASE HOMOLOG, ZNSSP6
; FILE REFERENCE: 98-80
; CURRENT APPLICATION NUMBER: US/09/482,180A
; CURRENT FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/115,721
; PRIOR FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1420
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (135)...(1271)
; US-09-482-180A-1

Query Match      8.2%; Score 48; DB 3; Length 1420;
Best Local Similarity 44.7%; Pred. No. 3.6e-05;
Matches 186; Conservative 0; Mismatches 230; Indels 0; Gaps 0;

Qy 3 GACATATCCAGAGGATTTCTAGAGCTCTATTACAACTCTGACCTGGAAGCATTGATG 62
Db 572 GACATCTCCAGTGGGACTTCACTGAGGACTTCTTCAACCTGACGCTCAAGGAGCTGCAC 731
Qy 63 GGCATAGAAATGGTCCATCGCTTTTGCTCAGGGGGGTTTGATGAAACAGACTCA 122
Db 732 CTGCAGCGCTGGGTGGTGGCTGCCTGCCCGCCAGGCCCAATTCATGCTAAAGGAGATGAC 791
Qy 123 GACATGTTTCATCAATGTTGACTATCTGACTGAACCTGCTTCTGAAGAAAAACAGAACACC 182
Db 792 GATGCTTTGTCACGCTCCCAACGTTGATAGTTCTTGGATGGTGGGACCCAGCCGAC 851
Qy 183 AGGTTTTTCACTGGCTTCTTGAACCTCAATGAGTTTCCCATCAGCAGGCCATTTCAGCAAG 242
Db 854 GACCTCTCGTGGGAGATGTATCGCCCAAGCGCTGCCCAACAGAACACACTAAGGTCAA 911
Qy 243 TGGTTTGTAGTAATCTGAATATCCGTGGGACAGGTACCCACCATTCGTCTCGGGCACC 302
Db 912 TACTTCATCCCAACCCCTCAATGTACAGGGCCACCCCACTACCCACCCCTATGCTGGTGGGGGA 971
Qy 303 GGCTACGTTGTTTCTGGCGAGTGGCGAGTCAGGTGTACAATGTCTCCAAGAGCGTCCCA 362
Db 972 GGATATGTCATGTCAGAGCCACAGTGGCGGCTCCAGGCTATCATGGAAGATGCTGAA 1031
Qy 363 TACATTAACCTGGAAGACGTGTTTGTGGGGCTCTGCCTCGAAAGGCTGAACATCAG 418
Db 1032 CTCCTCTCCATTGATGATGCTCTTGTGGGTATGTCCTGAGGAGGCTGGGGCTGAG 1087

RESULT 7
US-09-949-016-14938/c
; Sequence 14938, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13715
; LENGTH: 125902
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(125902)
; OTHER INFORMATION: n = A,T,C or G

; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell C.
; APPLICANT: Yamamoto, Gayle
; APPLICANT: Gao, Zeren
; APPLICANT: Whitmore, Stephen
; APPLICANT: Jaspers, Stephen
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING
; TITLE OF INVENTION: BETA-1,3-GALACTOSYLTRANSFERASE HOMOLOG, ZNSSP6
; FILE REFERENCE: 98-80
; CURRENT APPLICATION NUMBER: US/09/482,180A
; CURRENT FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/115,721
; PRIOR FILING DATE: 1999-01-12
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1420
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (135)...(1271)
; US-09-482-180A-1

Query Match      8.2%; Score 48; DB 4; Length 22374;
Best Local Similarity 44.7%; Pred. No. 0.00019;
Matches 186; Conservative 0; Mismatches 230; Indels 0; Gaps 0;

Qy 3 GACATATCCAGAGGATTTCTAGAGCTCTATTACAACTCTGACCTGGAAGCATTGATG 62
Db 21354 GACATCTCCAGTGGGACTTCACTGAGGACTTCTTCAACCTGACGCTCAAGGAGCTGCAC 21295
Qy 63 GGCATAGAAATGGTCCATCGCTTTTGCTCAGGGGGGTTTGATGAAACAGACTCA 122
Db 21294 CTGCAGCGCTGGGTGGTGGCTGCCTGCCCGCCAGGCCCAATTCATGCTAAAGGAGATGAC 21235
Qy 123 GACATGTTTCATCAATGTTGACTATCTGACTGAACCTGCTTCTGAAGAAAAACAGAACACC 182
Db 21234 GATGCTTTGTCACGCTCCCAACGTTGATAGTTCTTGGATGGTGGGACCCAGCCGAC 21175
Qy 183 AGGTTTTTCACTGGCTTCTTGAACCTCAATGAGTTTCCCATCAGCAGGCCATTTCAGCAAG 242
Db 21174 GACCTCTCGTGGGAGATGTATCGCCCAAGCGCTGCCCAACAGAACACACTAAGGTCAA 21115
Qy 243 TGGTTTGTAGTAATCTGAATATCCGTGGGACAGGTACCCACCATTCGTCTCGGGCACC 302
Db 21114 TACTTCATCCCAACCCCTCAATGTACAGGGCCACCCCACTACCCACCCCTATGCTGGTGGGGGA 21055
Qy 303 GGCTACGTTGTTTCTGGCGAGTGGCGAGTCAGGTGTACAATGTCTCCAAGAGCGTCCCA 362
Db 21054 GGATATGTCATGTCAGAGCCACAGTGGCGGCTCCAGGCTATCATGGAAGATGCTGAA 20995
Qy 363 TACATTAACCTGGAAGACGTGTTTGTGGGGCTCTGCCTCGAAAGGCTGAACATCAG 418
Db 20994 CTCITCCCAATTGATGATGCTCTTGTGGGTATGTCCTGAGGAGGCTGGGGCTGAG 20939

RESULT 8
US-09-949-016-13715/c
; Sequence 13715, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13715
; LENGTH: 125902
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(125902)
; OTHER INFORMATION: n = A,T,C or G
```


US-09-949-016-13715

Query Match 7.9%; Score 46; DB 4; Length 125902;
 Best Local Similarity 52.0%; Pred. No. 0.0026;
 Matches 103; Conservative 0; Mismatches 95; Indels 0; Gaps 0;
 QY 29 CGTCTATTACATCTGACCTGAAGACCATGATGGGCTAGATGGGTCATCGCTTTTG 88
 Db 9845 CGTCTCAACAAACAAACAAACAAACCTATTATGGCATTTGAGGTGGTAACTGAGTTTGG 9786
 QY 89 TCCTCAGCGGGGTTTGTGATGAAACAGACTCAGACATGTTTCATCAATGTTGACTATCT 148
 Db 9785 CCCCATTCGCACTACATCATGACAGACACCTATGTTGTTTCATCAATGAGCAATTT 9726
 QY 149 GACTGAACCTGCTTCTGAAGAAACAGAACACCACTGTTTTCATCTGGCTTCTTGAAACT 208
 Db 9725 AGTGTAGTATCTTTTAAACTTAAACCACTCAGAGAAGTTTTTTCACAGGTATCCCTAAT 9666
 QY 209 CAATGAGTTTCCCATCAG 226
 Db 9665 TGATAAATTATCCCATAG 9648

RESULT 9

US-09-055-097-2
 ; Sequence 2, Application US/09055097
 ; Patent No. 5955282
 ; GENERAL INFORMATION:
 ; APPLICANT: Hillman, Jennifer L.
 ; APPLICANT: Guegler, Karl J.
 ; APPLICANT: Corley, Neil C.
 ; APPLICANT: Shah, Purvi
 ; APPLICANT: Patterson, Chandra
 ; TITLE OF INVENTION: HUMAN OXIDIZED LDL RECEPTOR
 ; NUMBER OF SEQUENCES: 7
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSEQ for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/055,097
 ; FILING DATE: Filed Herewith
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Cerrone, Michael C.
 ; REGISTRATION NUMBER: 39,132
 ; REFERENCE/DOCKET NUMBER: PF-0490 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (650) 855-0555
 ; TELEFAX: (650) 845-4166
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1434 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: DUODNOT02
 ; CLONE: 1705085

Query Match 6.9%; Score 40.2; DB 2; Length 1434;

Best Local Similarity 54.4%; Pred. No. 0.014;
 Matches 81; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
 QY 1 GGGACATTATCCAGAAGGATTTCCTAGACGCTATTACAAATCTGACCTCGAAGACCATGA 60
 Db 446 GGGATACTTTCAGCGCCCTTCCAGGACTCTTACCGCACTCACCCTAAAGACCTCA 505
 QY 61 TGGGCATAGAAATGGTCCATCGCTTTTGTCTCAGGCGCGTTTGTGATGAAAAACAGACT 120
 Db 506 CGCGGCTGAACTGGGCTGAGAAACACTGCCCATGTGCCCGATCGTCTCAAGACGGACG 565
 QY 121 CAGACATGTTTCATCAATGTTGACTATCTG 149
 Db 566 ATGATGTGTATGTCAACGTCCTCGAATG 594

RESULT 10

US-09-373-902-2
 ; Sequence 2, Application US/09373902
 ; Patent No. 6649737
 ; GENERAL INFORMATION:
 ; APPLICANT: Hillman, Jennifer L.
 ; APPLICANT: Guegler, Karl J.
 ; APPLICANT: Corley, Neil C.
 ; APPLICANT: Shah, Purvi
 ; APPLICANT: Patterson, Chandra
 ; TITLE OF INVENTION: HUMAN OXIDIZED LDL RECEPTOR
 ; NUMBER OF SEQUENCES: 7
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSEQ for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/373,902
 ; FILING DATE: 12-Aug-1999
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/055,097
 ; FILING DATE: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Cerrone, Michael C.
 ; REGISTRATION NUMBER: 39,132
 ; REFERENCE/DOCKET NUMBER: PF-0490 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (650) 855-0555
 ; TELEFAX: (650) 845-4166
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1434 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: DUODNOT02
 ; CLONE: 1705085
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Query Match 6.9%; Score 40.2; DB 4; Length 1434;

Best Local Similarity 54.4%; Pred. No. 0.014;
 Matches 81; Conservative 0; Mismatches 68; Indels 0; Gaps 0;
 QY 1 GGGACATTATCCAGAAGGATTTCCTAGACGCTATTACAAATCTGACCTCGAAGACCATGA 60
 Db 446 GGGATACTTTCAGCGCCCTTCCAGGACTCTTACCGCACTCACCCTAAAGACCTCA 505

```
Qy 61 TGGGCATAGATGGTCCATCGCTTTTGTCTCCTCAGGGCGCTTTTGTGATGAACAGACT 120
Db 506 GCGGCTGAATGGGCTGAGAAACACTGCCCATGGCCCGATACGTCTCCTCAAGACGGACG 565
Qy 121 CAGACATGTTTCATCAATGTTGACTATCTG 149
Db 566 ATGATGTGTATGTCACAGCTCCTGACTG 594

RESULT 11
US-09-949-016-5696
; Sequence 5696, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5696
; LENGTH: 1640
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-5696

Query Match 6.9%; Score 40.2; DB 4; Length 1640;
Best Local Similarity 54.4%; Pred. No. 0.015; Indels 0; Gaps 0;
Matches 81; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

Qy 1 GGGACATTTATCCAGAAGGATTTCTTAGACGTCTTATTAACAATCTGACCTGAAGACCAATGA 60
Db 838 GGGATATCTTGCAGGCGCCCTTCCAGGACTCTACCGCAACTCACCCTAAAGACCCCTCA 697
Qy 61 TGGGCATAGATGGTCCATCGCTTTTGTCTCCTCAGGGCGCTTTTGTGATGAACAGACT 120
Db 598 GCGGCTGAATGGGCTGAGAAACACTGCCCATGGCCCGATACGTCTCCTCAAGACGGACG 757
Qy 121 CAGACATGTTTCATCAATGTTGACTATCTG 149
Db 758 ATGATGTGTATGTCACAGCTCCTGACTG 786

RESULT 12
US-09-949-016-17438
; Sequence 17438, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17438
; LENGTH: 5640
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17438

Query Match 6.9%; Score 40.2; DB 4; Length 8436;
Best Local Similarity 54.4%; Pred. No. 0.042; Indels 0; Gaps 0;
Matches 81; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

Qy 1 GGGACATTTATCCAGAAGGATTTCTTAGACGTCTTATTAACAATCTGACCTGAAGACCAATGA 60
Db 7758 GGGATATCTTGCAGGCGCCCTTCCAGGACTCTACCGCAACTCACCCTAAAGACCCCTCA 7817
Qy 61 TGGGCATAGATGGTCCATCGCTTTTGTCTCCTCAGGGCGCTTTTGTGATGAACAGACT 120
Db 7818 GCGGCTGAATGGGCTGAGAAACACTGCCCATGGCCCGATACGTCTCCTCAAGACGGACG 7877
Qy 121 CAGACATGTTTCATCAATGTTGACTATCTG 149
Db 7878 ATGATGTGTATGTCACAGCTCCTGACTG 7906

RESULT 14
US-09-270-767-444
; Sequence 444, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
```

```
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17438

Query Match 6.9%; Score 40.2; DB 4; Length 5640;
Best Local Similarity 54.4%; Pred. No. 0.033; Indels 0; Gaps 0;
Matches 81; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

Qy 1 GGGACATTTATCCAGAAGGATTTCTTAGACGTCTTATTAACAATCTGACCTGAAGACCAATGA 60
Db 2638 GGGATATCTTGCAGGCGCCCTTCCAGGACTCTACCGCAACTCACCCTAAAGACCCCTCA 2697
Qy 61 TGGGCATAGATGGTCCATCGCTTTTGTCTCCTCAGGGCGCTTTTGTGATGAACAGACT 120
Db 2698 GCGGCTGAATGGGCTGAGAAACACTGCCCATGGCCCGATACGTCTCCTCAAGACGGACG 2757
Qy 121 CAGACATGTTTCATCAATGTTGACTATCTG 149
Db 2758 ATGATGTGTATGTCACAGCTCCTGACTG 2786

RESULT 13
US-09-949-016-13963
; Sequence 13963, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13963
; LENGTH: 8436
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13963

Query Match 6.9%; Score 40.2; DB 4; Length 8436;
Best Local Similarity 54.4%; Pred. No. 0.042; Indels 0; Gaps 0;
Matches 81; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

Qy 1 GGGACATTTATCCAGAAGGATTTCTTAGACGTCTTATTAACAATCTGACCTGAAGACCAATGA 60
Db 7758 GGGATATCTTGCAGGCGCCCTTCCAGGACTCTACCGCAACTCACCCTAAAGACCCCTCA 7817
Qy 61 TGGGCATAGATGGTCCATCGCTTTTGTCTCCTCAGGGCGCTTTTGTGATGAACAGACT 120
Db 7818 GCGGCTGAATGGGCTGAGAAACACTGCCCATGGCCCGATACGTCTCCTCAAGACGGACG 7877
Qy 121 CAGACATGTTTCATCAATGTTGACTATCTG 149
Db 7878 ATGATGTGTATGTCACAGCTCCTGACTG 7906

RESULT 14
US-09-270-767-444
; Sequence 444, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
```

	Query Match	6.7%;	Score 39.2;	DB 4;	Length 748;
	Best Local Similarity	48.0%;	Pred. No. 0.02;		
	Matches 144;	Conservative 0;	Mismatches 153;	Indels 3;	Gaps 1;
Qy	86	TTGTCTTCAGCGCGGCTTGTGTGATGAAAAACAGACTCAGACATGTCATCAATCAATGTTGAC	145		
Dd	8	TTGCCGAAGGCAAGTACATTTCTCAGACGGACGACATGTTTCATCAATGTGCCCA	67		
Qy	146	TCTGACTGAATCTGCTTCTGTAAGAAAAACAGAACACCGAGTGTTCCTCGCTTCTTGAA	205		
Dd	68	GCTGCTGACCTACTGTGACAAAGCACAGGACAGCGTACCATATACGGTGTGTGGCCAA	127		
Qy	206	ACTCAATGAGTTTCCCATCAGGCAGCCATTCAGCAAGTGGTTTGTCAAGTAAATCTGAATA	265		
Dd	128	---GNAAGTGMAAACCGATTCGCAATTAAGAAATCCAAATACTATGTGTCCGTTCGATCAGTT	184		
Qy	266	TCCGTGGGACAGGTACCCACATTTCTGCTCCGGCACCGGCTACGTGTTTTCTCGCGACGT	325		
Dd	185	CGGGCGGAGTAATCCCGTCTTCCACTGGACCGCGCTACGTACTACCCGGAGACAT	244		
Qy	326	GGCGAGTCAGGTGTACAATGTTCTCCAAGACGCTCCCATACATTTAAATCTGGAACACGTGTT	385		

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 22, 2005, 14:50:12 ; Search time 1516.01 Seconds
(without alignments)
2575.655 Million cell updates/sec

Title: US-10-777-828-8_COPY_428_1011

Perfect score: 584

Sequence: 1 gggacattatccagaagat.....aagatttcgcgtgtctga 584

Scoring table:

IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 7400704 seqs, 3343079526 residues

Total number of hits satisfying chosen parameters: 14801408

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications NA:*
- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
 - 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
 - 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
 - 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
 - 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
 - 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
 - 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
 - 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
 - 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
 - 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
 - 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
 - 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
 - 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
 - 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
 - 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
 - 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
 - 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
 - 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
 - 19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
 - 20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
 - 21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
 - 22: /cgn2_6/ptodata/1/pubpna/US10J_NEW_PUB.seq.*
 - 23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
 - 24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
 - 25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
 - 26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	584	100.0	1011	19	US-10-777-828-8
2	570.4	97.7	681	13	Sequence 8, Appli
3	570.4	97.7	681	13	Sequence 21208, A
4	570.4	97.7	681	13	Sequence 21209, A
5	570.4	97.7	681	17	Sequence 21210, A
6	570.4	97.7	681	17	Sequence 21208, A
7	570.4	97.7	681	17	Sequence 21209, A
					Sequence 21210, A

8	438	75.0	516	13	US-10-027-632-106210	Sequence 106210,
9	438	75.0	516	17	US-10-027-632-106210	Sequence 106210,
10	437.6	74.9	516	13	US-10-027-632-106208	Sequence 106208,
11	437.6	74.9	516	13	US-10-027-632-106208	Sequence 106208,
12	437.6	74.9	516	17	US-10-027-632-106208	Sequence 106208,
13	437.6	74.9	516	17	US-10-027-632-106209	Sequence 106209,
14	119	20.4	1965	21	US-10-764-420-1500	Sequence 1500, Ap
15	114.4	19.6	1266	9	US-09-529-063-36	Sequence 26, Appl
16	114.4	19.6	1266	16	US-10-414-378-26	Sequence 1, Appl
17	114.4	19.6	1269	14	US-10-112-616A-1	Sequence 29, Appl
18	114.4	19.6	1269	21	US-10-764-420-29	Sequence 27, Appl
19	114.4	19.6	2420	9	US-09-529-063-27	Sequence 27, Appl
20	114.4	19.6	2420	16	US-10-414-378-27	Sequence 603, App
21	109	18.7	2168	21	US-10-887-553A-603	Sequence 481, App
22	106.2	18.2	1358	20	US-10-370-715B-481	Sequence 77, Appl
23	106.2	18.2	1358	21	US-10-651-237-77	Sequence 9, Appl
24	106.2	18.2	1358	21	US-10-782-413-77	Sequence 208, App
25	106.2	18.2	1773	9	US-09-739-451-9	Sequence 208, App
26	106.2	18.2	2095	9	US-09-989-722-208	Sequence 208, App
27	106.2	18.2	2095	9	US-09-989-723-208	Sequence 208, App
28	106.2	18.2	2095	9	US-09-989-279-208	Sequence 208, App
29	106.2	18.2	2095	9	US-09-989-727-208	Sequence 208, App
30	106.2	18.2	2095	9	US-09-989-731-208	Sequence 208, App
31	106.2	18.2	2095	9	US-09-989-732-208	Sequence 208, App
32	106.2	18.2	2095	9	US-09-991-073-208	Sequence 208, App
33	106.2	18.2	2095	9	US-09-990-442-208	Sequence 208, App
34	106.2	18.2	2095	9	US-09-991-163-208	Sequence 208, App
35	106.2	18.2	2095	9	US-09-993-604-208	Sequence 208, App
36	106.2	18.2	2095	9	US-09-990-456-208	Sequence 208, App
37	106.2	18.2	2095	9	US-09-989-721-208	Sequence 208, App
38	106.2	18.2	2095	9	US-09-992-598-208	Sequence 208, App
39	106.2	18.2	2095	9	US-09-989-293A-208	Sequence 208, App
40	106.2	18.2	2095	9	US-09-989-735-208	Sequence 208, App
41	106.2	18.2	2095	9	US-09-990-444-208	Sequence 208, App
42	106.2	18.2	2095	9	US-09-991-181-208	Sequence 208, App
43	106.2	18.2	2095	9	US-09-989-730-208	Sequence 208, App
44	106.2	18.2	2095	9	US-09-990-436-208	Sequence 208, App
45	106.2	18.2	2095	9	US-09-993-687-208	Sequence 208, App

ALIGNMENTS

RESULT 1

US-10-777-828-8
; Sequence 8, Application US/10777828
; Publication No. US20040142425A1
; GENERAL INFORMATION:
; APPLICANT: Clausen, Henrik
; APPLICANT: Amado, Margarita
; TITLE OF INVENTION: UDP-GALACTOSE: BETA-N-ACETYL-GLUCOSAMINE BETA 1,3
; TITLE OF INVENTION: GALACTOSYLTRANSFERASES, BETAGAL-T5
; FILE REFERENCE: 7188-157
; CURRENT APPLICATION NUMBER: US/10/777,828
; CURRENT FILING DATE: 2004-02-12
; PRIOR APPLICATION NUMBER: US/09/831,630
; PRIOR FILING DATE: 2001-05-10
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 1011
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (79)..(1008)
US-10-777-828-8

Query Match 100.0%; Score 584; DB 19; Length 1011;

Best Local Similarity 100.0%; Pred. No. 8.2e-188; Indels 0; Gaps 0;

Matches 584; Conservative 0; Mismatches 0;

Qy 1 GGGACATTATCCAGAGGATTTCTAGACGCTCTATTACAATCTGACCCCTGAAGACCATGA 60

```
Db 428 GGCACATATCCAGAGAGATTTCCTAGACGCTATTACAACTGACCCCTGAAGACCATGA 487
Qy 61 TGGGCATAGAAATGGGTCCATCGCTTTTCTCCTCAGGCGCGCTTTGTGATGAAACACAGACT 120
Db 488 TGGGCATAGAAATGGGTCCATCGCTTTTCTCCTCAGGCGCGCTTTGTGATGAAACACAGACT 547
Qy 121 CAGACATGTTTCATCAATCTGACTATCTGACTGACTGACTGACTGACTGACTGACTGACTGACT 180
Db 548 CAGACATGTTTCATCAATCTGACTATCTGACTGACTGACTGACTGACTGACTGACTGACTGACT 607
Qy 181 CCAGGTTTTTCACTGGCTTCTTTGAAACTCAATAGTTCCTCCATCAGGCGAGCCATTTCAGCA 240
Db 608 CCAGGTTTTTCACTGGCTTCTTTGAAACTCAATAGTTCCTCCATCAGGCGAGCCATTTCAGCA 667
Qy 241 AGTGGTTTGTGAGTAAATCTGAATATCCGTGGGACAGGTACCCACCAATTCCTGCTCCGGCA 300
Db 668 AGTGGTTTGTGAGTAAATCTGAATATCCGTGGGACAGGTACCCACCAATTCCTGCTCCGGCA 727
Qy 301 CCGGCTACGTGTTTCTCGGACGCTGGCGAGTCAGGTGTACAAATGTCTCCTCAAGAGCGTCC 360
Db 728 CCGGCTACGTGTTTCTCGGACGCTGGCGAGTCAGGTGTACAAATGTCTCCTCAAGAGCGTCC 787
Qy 361 CATACATTAAACTGGAAGACGCTGTTTGTGGGCTCTGCTCTCGAAAGGCTGAACATCAGAT 420
Db 788 CATACATTAAACTGGAAGACGCTGTTTGTGGGCTCTGCTCTCGAAAGGCTGAACATCAGAT 847
Qy 421 TGGAGGAGCTCCACTCCAGCGACCTTTTTTTCAGGGGGCTTACGCTTCTCCGTAATGCC 480
Db 848 TGGAGGAGCTCCACTCCAGCGACCTTTTTTTCAGGGGGCTTACGCTTCTCCGTAATGCC 907
Qy 481 TCCTCAGAGAGATCGTGCGCTGCCACTTCATCAGCCCTCGGACTCTCTTGACTACTGGC 540
Db 908 TCCTCAGAGAGATCGTGCGCTGCCACTTCATCAAGCCCTCGGACTCTCTTGACTACTGGC 967
Qy 541 AGGCTCTAGAGAAATCCCGGGGGGAGAAATGTCCGCTGTCTGA 584
Db 968 AGGCTCTAGAGAAATCCCGGGGGGAGAAATGTCCGCTGTCTGA 1011
```

RESULT 2

```
US-10-027-632-21208
; Sequence 21208, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21208
; LENGTH: 681
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-21208
```

```
Query Match 97.7% Score 570.4; DB 13; Length 681;
Best Local Similarity 99.1%; Pred. No. 2.9e-183;
Matches 571; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
```

```
Qy 1 GGGACATTATCCAGAGAGATTTCCTAGACGCTATTACAACTGACCCCTGAAGACCATGA 60
Db 76 GGGACATTATCCAGAGAGATTTCCTAGACGCTATTACAACTGACCCCTGAAGACCATGA 135
Qy 61 TGGGCATAGAAATGGGTCCATCGCTTTTCTCCTCAGGCGCGCTTTGTGATGAAACACAGACT 120
Db 136 TGGGCATAGAAATGGGTCCATCGCTTTTCTCCTCAGGCGCGCTTTGTGATGAAACACAGACT 195
Qy 121 CAGACATGTTTCATCAATCTGACTATCTGACTGACTGACTGACTGACTGACTGACTGACT 180
Db 196 CAGACATGTTTCATCAATCTGACTATCTGACTGACTGACTGACTGACTGACTGACTGACT 255
Qy 181 CCAGGTTTTTCACTGGCTTCTTTGAAACTCAATAGTTCCTCCATCAGGCGAGCCATTTCAGCA 240
Db 256 CCAGGTTTTTCACTGGCTTCTTTGAAACTCAATAGTTCCTCCATCAGGCGAGCCATTTCAGCA 315
Qy 241 AGTGGTTTGTGAGTAAATCTGAATATCCGTGGGACAGGTACCCACCAATTCCTGCTCCGGCA 300
Db 316 AGTGGTTTGTGAGTAAATCTGAATATCCGTGGGACAGGTACCCACCAATTCCTGCTCCGGCA 375
Qy 301 CCGGCTACGTGTTTCTCGGACGCTGGCGAGTCAGGTGTACAAATGTCTCCTCAAGAGCGTCC 360
Db 376 CCGGCTACGTGTTTCTCGGACGCTGGCGAGTCAGGTGTACAAATGTCTCCTCAAGAGCGTCC 435
Qy 361 CATACATTAAACTGGAAGACGCTGTTTGTGGGCTCTGCTCTCGAAAGGCTGAACATCAGAT 420
Db 436 CATACATTAAACTGGAAGACGCTGTTTGTGGGCTCTGCTCTCGAAAGGCTGAACATCAGAT 495
Qy 421 TGGAGGAGCTCCACTCCAGCGACCTTTTTTTCAGGGGGCTTACGCTTCTCCGTAATGCC 480
Db 496 TGGAGGAGCTCCACTCCAGCGACCTTTTTTTCAGGGGGCTTACGCTTCTCCGTAATGCC 555
Qy 481 TCCTCAGAGAGATCGTGCGCTGCCACTTCATCAGCCCTCGGACTCTCTTGACTACTGGC 540
Db 556 TCCTCAGAGAGATCGTGCGCTGCCACTTCATCAAGCCCTCGGACTCTCTTGACTACTGGC 615
Qy 541 AGGCTCTAGAGAAATCCCGGGGGGAGAAATGTTCGCGC 576
Db 616 AGGCTCTAGAGAAATCCCGGGGGGAGAAATGTTCGCGC 651
```

RESULT 3

```
US-10-027-632-21209
; Sequence 21209, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21209
```

```
; LENGTH: 681
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-21209

Query Match          97.7%; Score 570.4; DB 13; Length 681;
Best Local Similarity 99.1%; Pred. No. 2.9e-183;
Matches 571; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGGACATTATCCAGAAAGGATTTCTTAGACGCTCTATTACAATCTGACCCCTGAAGACCATGA 60
Db 76 GGGACATTATCCAGAAAGGATTTCTTAGACGCTCTATTACAATCTGACCCCTGAAGACCATGA 135
QY 61 TGGGCATAGAATGGGTCCATCGCTTTTGTCTCAGCGCGGCTTTGTGATGAATAACAGACT 120
Db 136 TGGGCATAGAATGGGTCCATCGCTTTTGTCTCAGCGCGGCTTTGTGATGAATAACAGACT 195
QY 121 CAGACATGTTCAATCAATGTTGACTATCTGACTGAACTGCTCTGAAGAAACAGAA 180
Db 196 CAGACATGTTCAATCAATGTTGACTATCTGACTGAACTGCTCTGAAGAAACAGAA 255
QY 181 CCAGGTTTTTCACTGGCTTCTTGAAGTCTTGAAGTCTTCCATCAGGCGAGCCATTTCAGCA 240
Db 256 CCAGGTTTTTCACTGGCTTCTTGAAGTCTTGAAGTCTTCCATCAGGCGAGCCATTTCAGCA 315
QY 241 AGTGGTTTGTCAAGTAAATCTGAATATCCGTGGGACAGGTACCACCAATTCGTCTCCGGCA 300
Db 316 AGTGGTTTGTCAAGTAAATCTGAATATCCGTGGGACAGGTACCACCAATTCGTCTCCGGCA 375
QY 301 CCGGCTACGTGTTTCTGGCAGCTGGGAGTCCAGGTGACAACTGCTCCAGAGCGTCC 360
Db 376 CCGGCTACGTGTTTCTGGCAGCTGGGAGTCCAGGTGACAACTGCTCCAGAGCGTCC 435
QY 361 CATACATTAAACTGGAAGACGTTGTTGGGGCTCTGCTCGAAAGGCTGAACATCAGAT 420
Db 436 CATACATTAAACTGGAAGACGTTGTTGGGGCTCTGCTCGAAAGGCTGAACATCAGAT 495
QY 421 TGGAGGAGCTCCACTCCAGCGGACCTTTTTCAGGGGGCTTACGCTTCTCCGATGCC 480
Db 496 TGGAGGAGCTCCACTCCAGCGGACCTTTTTCAGGGGGCTTACGCTTCTCCGATGCC 555
QY 481 TCTTCAGAGGATCGTGCGCTGCCATTCATCAAGCCTCGGACTCTCTTGGACTACTGGC 540
Db 556 TCTTCAGAGGATCGTGCGCTGCCATTCATCAAGCCTCGGACTCTCTTGGACTACTGGC 615
QY 541 AGGCTCTAGAGAAATCCCGGGGGGAAGATTGTCCGC 576
Db 616 AGGCTCTAGAGAAATCCCGGGGGGAAGATTGTCCGC 651

RESULT 4
US-10-027-632-21210
; Sequence 21210, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 10827,129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
```

```
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21210
; LENGTH: 681
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-21210

Query Match          97.7%; Score 570.4; DB 13; Length 681;
Best Local Similarity 99.1%; Pred. No. 2.9e-183;
Matches 571; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GGGACATTATCCAGAAAGGATTTCTTAGACGCTCTATTACAATCTGACCCCTGAAGACCATGA 60
Db 76 GGGACATTATCCAGAAAGGATTTCTTAGACGCTCTATTACAATCTGACCCCTGAAGACCATGA 135
QY 61 TGGGCATAGAATGGGTCCATCGCTTTTGTCTCAGCGCGGCTTTGTGATGAATAACAGACT 120
Db 136 TGGGCATAGAATGGGTCCATCGCTTTTGTCTCAGCGCGGCTTTGTGATGAATAACAGACT 195
QY 121 CAGACATGTTCAATCAATGTTGACTATCTGACTGAACTGCTCTGAAGAAACAGAA 180
Db 196 CAGACATGTTCAATCAATGTTGACTATCTGACTGAACTGCTCTGAAGAAACAGAA 255
QY 181 CCAGGTTTTTCACTGGCTTCTTGAAGTCTTGAAGTCTTCCATCAGGCGAGCCATTTCAGCA 240
Db 256 CCAGGTTTTTCACTGGCTTCTTGAAGTCTTGAAGTCTTCCATCAGGCGAGCCATTTCAGCA 315
QY 241 AGTGGTTTGTCAAGTAAATCTGAATATCCGTGGGACAGGTACCACCAATTCGTCTCCGGCA 300
Db 316 AGTGGTTTGTCAAGTAAATCTGAATATCCGTGGGACAGGTACCACCAATTCGTCTCCGGCA 375
QY 301 CCGGCTACGTGTTTCTGGCAGCTGGGAGTCCAGGTGACAACTGCTCCAGAGCGTCC 360
Db 376 CCGGCTACGTGTTTCTGGCAGCTGGGAGTCCAGGTGACAACTGCTCCAGAGCGTCC 435
QY 361 CATACATTAAACTGGAAGACGTTGTTGGGGCTCTGCTCGAAAGGCTGAACATCAGAT 420
Db 436 CATACATTAAACTGGAAGACGTTGTTGGGGCTCTGCTCGAAAGGCTGAACATCAGAT 495
QY 421 TGGAGGAGCTCCACTCCAGCGGACCTTTTTCAGGGGGCTTACGCTTCTCCGATGCC 480
Db 496 TGGAGGAGCTCCACTCCAGCGGACCTTTTTCAGGGGGCTTACGCTTCTCCGATGCC 555
QY 481 TCTTCAGAGGATCGTGCGCTGCCATTCATCAAGCCTCGGACTCTCTTGGACTACTGGC 540
Db 556 TCTTCAGAGGATCGTGCGCTGCCATTCATCAAGCCTCGGACTCTCTTGGACTACTGGC 615
QY 541 AGGCTCTAGAGAAATCCCGGGGGGAAGATTGTCCGC 576
Db 616 AGGCTCTAGAGAAATCCCGGGGGGAAGATTGTCCGC 651

RESULT 5
US-10-027-632-21208
; Sequence 21208, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 10827,129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
```


; FILE REFERENCE: 108827.129
 ; CURRENT APPLICATION NUMBER: US/10/027,632
 ; PENDING FILING DATE: 2002-04-30
 ; PRIOR APPLICATION NUMBER: US 60/218,006
 ; PRIOR FILING DATE: 2000-07-12
 ; PRIOR APPLICATION NUMBER: US 60/198,676
 ; PRIOR FILING DATE: 2000-04-20
 ; PRIOR APPLICATION NUMBER: US 60/193,483
 ; PRIOR FILING DATE: 2000-03-29
 ; PRIOR APPLICATION NUMBER: US 60/185,218
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: US 60/167,363
 ; PRIOR FILING DATE: 1999-11-23
 ; PRIOR APPLICATION NUMBER: US 60/156,358
 ; PRIOR FILING DATE: 1999-09-28
 ; PRIOR APPLICATION NUMBER: US 60/146,002
 ; PRIOR FILING DATE: 1999-08-09
 ; NUMBER OF SEQ ID NOS: 325720
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 21210
 ; LENGTH: 681
 ; TYPE: DNA
 ; ORGANISM: Human
 ; US-10-027-632-21210

Query Match 97.7%; Score 570.4; DB 17; Length 681;
 Best Local Similarity 99.1%; Pred. No. 2.9e-183;
 Matches 571; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 GGGACATTTCCAGAAAGATTCTTAGAGCTTATTACAACTGACCCCTGAAGACCATGA 60
 DB 76 GGGACATTTCCAGAAAGATTCTTAGAGCTTATTACAACTGACCCCTGAAGACCATGA 135
 QY 61 TGGGCATAGAAATGGGTCCATCGCTTTTCTCAGGCGCGCTTTGTGATGAAGACAGACT 120
 DB 136 TGGGCATAGAAATGGGTCCATCGCTTTTCTCAGGCGCGCTTTGTGATGAAGACAGACT 195
 QY 121 CAGACATGTTTCATCAATGTTGACTATCTGACTGAACTGCTTTCTGAAGAAAAACAGAA 180
 DB 196 CAGACATGTTTCATCAATGTTGACTATCTGACTGAACTGCTTTCTGAAGAAAAACAGAA 255
 QY 181 CAGGTTTTCATCGGCTTTCTGAAGAACTCAATGAGTTTCCATCAGGAGCCATTCAGCA 240
 DB 256 CAGGTTTTCATCGGCTTTCTGAAGAACTCAATGAGTTTCCATCAGGAGCCATTCAGCA 315
 QY 241 AGTGTGTTGTGAGTAAATCTGAATATCCGTGGGACAGGTACCCACTTCTGCTCCGGCA 300
 DB 316 AGTGTGTTGTGAGTAAATCTGAATATCCGTGGGACAGGTACCCACTTCTGCTCCGGCA 375
 QY 301 CCGGCTACGTGTTTCTGGGACAGTGGGAGTCAAGTGTACAAATGTTCTTCAAGAGCGTCC 360
 DB 376 CCGGCTACGTGTTTCTGGGACAGTGGGAGTCAAGTGTACAAATGTTCTTCAAGAGCGTCC 435
 QY 361 CATACATTAACATGAAGACGTGTTTGTGGGCTCTGCTCGAAGGCTGAACATCAGAT 420
 DB 436 CATACATTAACATGAAGACGTGTTTGTGGGCTCTGCTCGAAGGCTGAACATCAGAT 495
 QY 421 TGGAGGAGCTCCACTCCAGCGGACCTTTTTCAGAGGGGCTTACGCTTCTCCGTATGCC 480
 DB 496 TGGAGGAGCTCCACTCCAGCGGACCTTTTTCAGAGGGGCTTACGCTTCTCCGTATGCC 555
 QY 481 TCTTCAGGAGGATCGTGGGCTCGCACTTTCATCAAGCCCTCGGACTCTCTTGAGCTACTGCC 540
 DB 556 TCTTCAGGAGGATCGTGGGCTCGCACTTTCATCAAGCCCTCGGACTCTCTTGAGCTACTGCC 615
 QY 541 AGGCTCTAGAGAAATTCGCGGGGGAAGATTGTCGC 576
 DB 616 AGGCTCTAGAGAAATTCGCGGGGGAAGATTGTCGC 651

RESULT 8
 US-10-027-632-106210
 ; Sequence 106210, Application US/10027632
 ; Publication No. US20030204075A9
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.

; Publication No. US20020198371A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 ; Polymorphisms in the Human Genome
 ; FILE REFERENCE: 108827.129
 ; CURRENT APPLICATION NUMBER: US/10/027,632
 ; PENDING FILING DATE: 2002-04-30
 ; PRIOR APPLICATION NUMBER: US 60/218,006
 ; PRIOR FILING DATE: 2000-07-12
 ; PRIOR APPLICATION NUMBER: US 60/198,676
 ; PRIOR FILING DATE: 2000-04-20
 ; PRIOR APPLICATION NUMBER: US 60/193,483
 ; PRIOR FILING DATE: 2000-03-29
 ; PRIOR APPLICATION NUMBER: US 60/185,218
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: US 60/167,363
 ; PRIOR FILING DATE: 1999-11-23
 ; PRIOR APPLICATION NUMBER: US 60/156,358
 ; PRIOR FILING DATE: 1999-09-28
 ; PRIOR APPLICATION NUMBER: US 60/146,002
 ; PRIOR FILING DATE: 1999-08-09
 ; NUMBER OF SEQ ID NOS: 325720
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 106210
 ; LENGTH: 516
 ; TYPE: DNA
 ; ORGANISM: Human
 ; US-10-027-632-106210

Query Match 75.0%; Score 438; DB 13; Length 516;
 Best Local Similarity 100.0%; Pred. No. 3.5e-138;
 Matches 438; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGGACATTTCCAGAAAGATTCTTAGAGCTTATTACAACTGACCCCTGAAGACCATGA 60
 DB 79 GGGACATTTCCAGAAAGATTCTTAGAGCTTATTACAACTGACCCCTGAAGACCATGA 138
 QY 61 TGGGCATAGAAATGGGTCCATCGCTTTTCTCAGGCGCGCTTTGTGATGAAGACAGACT 120
 DB 139 TGGGCATAGAAATGGGTCCATCGCTTTTCTCAGGCGCGCTTTGTGATGAAGACAGACT 198
 QY 121 CAGACATGTTTCATCAATGTTGACTATCTGACTGAACTGCTTTCTGAAGAAAAACAGAA 180
 DB 199 CAGACATGTTTCATCAATGTTGACTATCTGACTGAACTGCTTTCTGAAGAAAAACAGAA 258
 QY 181 CAGGTTTTCATCGGCTTTCTGAAGAACTCAATGAGTTTCCATCAGGAGCCATTCAGCA 240
 DB 259 CAGGTTTTCATCGGCTTTCTGAAGAACTCAATGAGTTTCCATCAGGAGCCATTCAGCA 318
 QY 241 AGTGTGTTGTGAGTAAATCTGAATATCCGTGGGACAGGTACCCACTTCTGCTCCGGCA 300
 DB 319 AGTGTGTTGTGAGTAAATCTGAATATCCGTGGGACAGGTACCCACTTCTGCTCCGGCA 378
 QY 301 CCGGCTACGTGTTTCTGGGACAGTGGGAGTCAAGTGTACAAATGTTCTTCAAGAGCGTCC 360
 DB 379 CCGGCTACGTGTTTCTGGGACAGTGGGAGTCAAGTGTACAAATGTTCTTCAAGAGCGTCC 438
 QY 361 CATACATTAACATGAAGACGTGTTTGTGGGCTCTGCTCGAAGGCTGAACATCAGAT 420
 DB 439 CATACATTAACATGAAGACGTGTTTGTGGGCTCTGCTCGAAGGCTGAACATCAGAT 498
 QY 421 TGGAGGAGCTCCACTCC 438
 DB 499 TGGAGGAGCTCCACTCC 516

RESULT 9
 US-10-027-632-106210
 ; Sequence 106210, Application US/10027632
 ; Publication No. US20030204075A9
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.

```
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 106210
; LENGTH: 516
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-106210
```

```
Query Match 75.0%; Score 438; DB 17; Length 516;
Best Local Similarity 100.0%; Pred. No. 3.5e-138;
Matches 438; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGACATTATCCAGAAGGATTTCTAGACGCTATTACAACTGACCCCTGAAGACCATGA 60
Db 79 GGGACATTATCCAGAAGGATTTCTAGACGCTATTACAACTGACCCCTGAAGACCATGA 138

Qy 61 TGGGCATAGAAATGGGTCCATCGCTTTTCTCCTCAGCGCGCGTGTGTGATGAAACAGACT 120
Db 139 TGGGCATAGAAATGGGTCCATCGCTTTTCTCCTCAGCGCGCGTGTGTGATGAAACAGACT 198

Qy 121 CAGACATGTTCAATGTTGACTATCTGACTGAACTCAATGAGTTTCCATCAGGAGCAACAA 180
Db 199 CAGACATGTTCAATGTTGACTATCTGACTGAACTCAATGAGTTTCCATCAGGAGCAACAA 258

Qy 181 CCAGGTTTTTCACTGGCTTCTTGAACCTCAATGAGTTTCCATCAGGAGCAACAA 240
Db 259 CCAGGTTTTTCACTGGCTTCTTGAACCTCAATGAGTTTCCATCAGGAGCAACAA 318

Qy 241 AGTGGTTTGTTCAGTAAATCTGAATATCCGTGGGACAGGTACCCACCAATTTCTGCTCCGGCA 300
Db 319 AGTGGTTTGTTCAGTAAATCTGAATATCCGTGGGACAGGTACCCACCAATTTCTGCTCCGGCA 378

Qy 301 CCGGCTACGTGTTTCTGGGACAGTGGGAGTCAAGTGTACAATGTCTCCAAAGAGCGTCC 360
Db 379 CCGGCTACGTGTTTCTGGGACAGTGGGAGTCAAGTGTACAATGTCTCCAAAGAGCGTCC 438

Qy 361 CATACATTAACCTGAAGACGCTGTTGTGGGGCTCTGCCCTCGAAAGGCTGAACATCAGAT 420
Db 439 CATACATTAACCTGAAGACGCTGTTGTGGGGCTCTGCCCTCGAAAGGCTGAACATCAGAT 498

Qy 421 TGGAGGAGCTCCACTCCC 438
Db 499 TGGAGGAGCTCCACTCCC 516
```

```
RESULT 10
US-10-027-632-106208
; Sequence 106208, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
```

```
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 106208
; LENGTH: 516
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-106208
```

```
Query Match 74.9%; Score 437.6; DB 13; Length 516;
Best Local Similarity 99.8%; Pred. No. 4.8e-138;
Matches 437; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGACATTATCCAGAAGGATTTCTAGACGCTATTACAACTGACCCCTGAAGACCATGA 60
Db 79 GGGACATTATCCAGAAGGATTTCTAGACGCTATTACAACTGACCCCTGAAGACCATGA 138

Qy 61 TGGGCATAGAAATGGGTCCATCGCTTTTCTCCTCAGCGCGCGTGTGTGATGAAACAGACT 120
Db 139 TGGGCATAGAAATGGGTCCATCGCTTTTCTCCTCAGCGCGCGTGTGTGATGAAACAGACT 198

Qy 121 CAGACATGTTCAATGTTGACTATCTGACTGAACTCAATGAGTTTCCATCAGGAGCAACAA 180
Db 199 CAGACATGTTCAATGTTGACTATCTGACTGAACTCAATGAGTTTCCATCAGGAGCAACAA 258

Qy 181 CCAGGTTTTTCACTGGCTTCTTGAACCTCAATGAGTTTCCATCAGGAGCAACAA 240
Db 259 CCAGGTTTTTCACTGGCTTCTTGAACCTCAATGAGTTTCCATCAGGAGCAACAA 318

Qy 241 AGTGGTTTGTTCAGTAAATCTGAATATCCGTGGGACAGGTACCCACCAATTTCTGCTCCGGCA 300
Db 319 AGTGGTTTGTTCAGTAAATCTGAATATCCGTGGGACAGGTACCCACCAATTTCTGCTCCGGCA 378

Qy 301 CCGGCTACGTGTTTCTGGGACAGTGGGAGTCAAGTGTACAATGTCTCCAAAGAGCGTCC 360
Db 379 CCGGCTACGTGTTTCTGGGACAGTGGGAGTCAAGTGTACAATGTCTCCAAAGAGCGTCC 438

Qy 361 CATACATTAACCTGAAGACGCTGTTGTGGGGCTCTGCCCTCGAAAGGCTGAACATCAGAT 420
Db 439 CATACATTAACCTGAAGACGCTGTTGTGGGGCTCTGCCCTCGAAAGGCTGAACATCAGAT 498

Qy 421 TGGAGGAGCTCCACTCCC 438
Db 499 TGGAGGAGCTCCACTCCC 516
```

```
RESULT 11
US-10-027-632-106209
; Sequence 106209, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
```

; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 106209
; LENGTH: 516
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-106209

Query Match 74.9%; Score 437.6; DB 13; Length 516;
Best Local Similarity 99.8%; Pred. No. 4.8e-138;
Matches 437; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGACATTATCCAGAAGGATTTCTAGAGCTCTATTACAATCTGACCCCTGAAGACCATGA 60
DB 79 GGGACATTATCCAGAAGGATTTCTAGAGCTCTATTACAATCTGACCCCTGAAGACCATGA 138
QY 61 TGGGCATAGATGGGTCCATCGCTTTTGTCTCAGCGCGGCTTTGTGATGAAAAACAGACT 120
DB 139 TGGGCATAGATGGGTCCATCGCTTTTGTCTCAGCGCGGCTTTGTGATGAAAAACAGACT 198
QY 121 CAGACATGTTCAATCAATGTTGACTATCTGACTGAATCTGCTCTGAAGAAAAACAGAA 180
DB 199 CAGACATGTTCAATCAATGTTGACTATCTGACTGAATCTGCTCTGAAGAAAAACAGAA 258
QY 181 CCAGGTTTTTCACTGGCTTTTGAAGAACTCAATGAGTTTCCCATCAGGAGCCATTTCAGCA 240
DB 259 CCAGGTTTTTCACTGGCTTTTGAAGAACTCAATGAGTTTCCCATCAGGAGCCATTTCAGCA 318
QY 241 AGTGGTTTGTGAGTAAATCTGAATATCGGTGGGACAGGTACCAATCTCTCCAGAGCGTCC 300
DB 319 AGTGGTTTGTGAGTAAATCTGAATATCGGTGGGACAGGTACCAATCTCTCCAGAGCGTCC 378
QY 301 CCGGCTACGTGTTTCTGCGGACGTGGGAGTCAGGTGTACAAATGTCTCCAGAGCGTCC 360
DB 379 CCGGCTACGTGTTTCTGCGGACGTGGGAGTCAGGTGTACAAATGTCTCCAGAGCGTCC 438
QY 361 CATACATTAAACTGGAAGACGTGTTTGTGGGGCTCTGCTCGAAAGGCTGAACATCAGAT 420
DB 439 CATACATTAAACTGGAAGACGTGTTTGTGGGGCTCTGCTCGAAAGGCTGAACATCAGAT 498
QY 421 TGGAGGAGCTCCACTCCC 438
DB 499 TGGAGGAGCTCCACTCCC 516

RESULT 12

US-10-027-632-106209
; Sequence 106208, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20

; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 106208
; LENGTH: 516
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-106208

Query Match 74.9%; Score 437.6; DB 17; Length 516;
Best Local Similarity 99.8%; Pred. No. 4.8e-138;
Matches 437; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGACATTATCCAGAAGGATTTCTAGAGCTCTATTACAATCTGACCCCTGAAGACCATGA 60
DB 79 GGGACATTATCCAGAAGGATTTCTAGAGCTCTATTACAATCTGACCCCTGAAGACCATGA 138
QY 61 TGGGCATAGATGGGTCCATCGCTTTTGTCTCAGCGCGGCTTTGTGATGAAAAACAGACT 120
DB 139 TGGGCATAGATGGGTCCATCGCTTTTGTCTCAGCGCGGCTTTGTGATGAAAAACAGACT 198
QY 121 CAGACATGTTCAATCAATGTTGACTATCTGACTGAATCTGCTCTGAAGAAAAACAGAA 180
DB 199 CAGACATGTTCAATCAATGTTGACTATCTGACTGAATCTGCTCTGAAGAAAAACAGAA 258
QY 181 CCAGGTTTTTCACTGGCTTTTGAAGAACTCAATGAGTTTCCCATCAGGAGCCATTTCAGCA 240
DB 259 CCAGGTTTTTCACTGGCTTTTGAAGAACTCAATGAGTTTCCCATCAGGAGCCATTTCAGCA 318
QY 241 AGTGGTTTGTGAGTAAATCTGAATATCGGTGGGACAGGTACCAATCTCTCCAGAGCGTCC 300
DB 319 AGTGGTTTGTGAGTAAATCTGAATATCGGTGGGACAGGTACCAATCTCTCCAGAGCGTCC 378
QY 301 CCGGCTACGTGTTTCTGCGGACGTGGGAGTCAGGTGTACAAATGTCTCCAGAGCGTCC 360
DB 379 CCGGCTACGTGTTTCTGCGGACGTGGGAGTCAGGTGTACAAATGTCTCCAGAGCGTCC 438
QY 361 CATACATTAAACTGGAAGACGTGTTTGTGGGGCTCTGCTCGAAAGGCTGAACATCAGAT 420
DB 439 CATACATTAAACTGGAAGACGTGTTTGTGGGGCTCTGCTCGAAAGGCTGAACATCAGAT 498
QY 421 TGGAGGAGCTCCACTCCC 438
DB 499 TGGAGGAGCTCCACTCCC 516

RESULT 13

US-10-027-632-106209
; Sequence 106209, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218

; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 106209
; LENGTH: 516
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-106209

Query Match 74.9%; Score 437.6; DB 17; Length 516;
Best Local Similarity 99.8%; Pred. No. 4.8e-138;
Matches 437; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGGACATTATCCAGAAGGATTTCTAGACGCTATTACAATCTGACCTGAGACCATGA 60
DB 75 GGGACATTATCCAGAAGGATTTCTAGACGCTATTACAATCTGACCTGAGACCATGA 138
QY 61 TGGGCATAGAATGGTCCATCGCTTTTCTCCTCAGGCGCGCTTTTGTGATGAAAACAGACT 120
DB 139 TGGGCATAGAATGGTCCATCGCTTTTCTCCTCAGGCGCGCTTTTGTGATGAAAACAGACT 198
QY 121 CAGACATGTTTCATCAATGTTGACTATCTGACTGAACTGCTCTTCTGAAGAAAACAGAA 180
DB 199 CAGACATGTTTCATCAATGTTGACTATCTGACTGAACTGCTCTTCTGAAGAAAACAGAA 258
QY 181 CCAGGTTTTTCACTGGCTCTTGTGAACCTCAATGATGTTTCCCATCAGGCGGCATTCAGCA 240
DB 259 CCAGGTTTTTCACTGGCTCTTGTGAACCTCAATGATGTTTCCCATCAGGCGGCATTCAGCA 318
QY 241 AGTGGTTGTGTCAGTAATCTGAATATCCGTGGGACAGGTACCCACCATTTCTGCTCCGGCA 300
DB 119 AGTGGTTGTGTCAGTAATCTGAATATCCGTGGGACAGGTACCCACCATTTCTGCTCCGGCA 378
QY 301 CCGGCTACGTTGTTTCTGGGACGCTGGGAGTCAAGTGTACAATGTTCTTCCAAAGAGCGTCC 360
DB 379 CCGGCTACGTTGTTTCTGGGACGCTGGGAGTCAAGTGTACAATGTTCTTCCAAAGAGCGTCC 438
QY 361 CATACATTAAACTGGAAGACCTGTTTGTGGGGCTCTGCTCGAAAGGCTGAACATCAGAT 420
DB 439 CATACATTAAACTGGAAGACCTGTTTGTGGGGCTCTGCTCGAAAGGCTGAACATCAGAT 498
QY 421 TGGAGGAGCTCCACTCCC 438
DB 439 TGGAGGAGCTCCACTCCC 516

RESULT 14
US-10-764-420-1500
; Sequence 150, Application US/10764420
; Publication No. US20050084872A1
; GENERAL INFORMATION:
; APPLICANT: Lum, Pek Yee
; APPLICANT: Tan, Yejun
; APPLICANT: Dai, Hongyue
; TITLE OF INVENTION: Methods For Determining Whether An Agent Possesses A Defined Biological Activity
; FILE REFERENCE: ROSA122057
; CURRENT APPLICATION NUMBER: US/10/764,420
; PRIOR FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US 60/442,797
; PRIOR FILING DATE: 2003-01-24
; PRIOR APPLICATION NUMBER: US 60/474,413
; PRIOR FILING DATE: 2003-05-30
; NUMBER OF SEQ ID NOS: 3683
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1500
; LENGTH: 1965

; TYPE: DNA
; ORGANISM: Mus musculus
US-10-764-420-1500
Query Match 20.4%; Score 119; DB 21; Length 1965;
Best Local Similarity 51.2%; Pred. No. 4.4e-29;
Matches 278; Conservative 0; Mismatches 265; Indels 0; Gaps 0;
QY 1 GGGACATTATCCAGAAGGATTTCTAGACGCTATTACAATCTGACCTGAGACCATGA 60
DB 646 GTGATATTATACGCAAGACTTTCTAGACATATAATAAATCTTGACCTTGAACACCATTA 705
QY 61 TGGGCATAGAATGGTCCATCGCTTTTGTCTCAGGCGCGCTTTGTGATGAAAACAGACT 120
DB 706 TGGCTTTTCAGTGGGTAATGGAGTTTGGCCCAATGCCAAGTATATATGAAACACACA 765
QY 121 CAGACATGTTTCATCAATGTTGACTATCTGACTGAACTGCTTCTGAAGAAAACAGAA 180
DB 766 CTGATGTTTTCATCAATGTTGCAATTTAGTCTATTATCTTTTAAACCTAAACCACTCAG 825
QY 181 CCAGGTTTTTCACTGGCTCTTGTGAACCTCAATGATGTTTCCCATCAGGCGGCATTCAGCA 240
DB 826 AGAAGTTTTTTCAGCGGCTATCTCTTAATGATAAATCTTATATAGAGGATTTTTCATA 885
QY 241 AGTGGTTGTGTCAGTAATCTGAATATCCGTGGGACAGGTACCCACCATTTCTGCTCCGGCA 300
DB 886 AAACACCATTTTCATCAAGAGTACCCCTTCAAGGTGTTCCCTCTCTGACAGCGGGC 945
QY 301 CCGGCTACGTTGTTTCTGGGACGCTGGCGAGTCAAGTGTACAATGTTCTCAAGAGCGTCC 360
DB 946 TGGGCTACATTATGTCGCGGCACTGGTGGCCAGGCTACGAGATGATGATGATCAGCTGA 1005
QY 361 CATACATTAAACTGGAAGACGCTGTTTGTGGGGCTCTGCTCGAAAGGCTGAACATCAGAT 420
DB 1006 AGCCCATCAAGTTTGAAGACGCTTTATGTTGGCATCTGTTTGAATTTGTTAAAGTGGA 1065
QY 421 TGGAGGAGCTCCACTCCCAGCGACCTTTTTCAGGGGGCTTACGCTTCTCCGCTATGCC 480
DB 1066 TTCTATTTCCAGAAGACACAACCTTTTCTTCTGTACAGATCCATTTGGATGATGTC 1125
QY 481 TCTTCAGAGAGATGTTGGCTGGCCTGCACTTCATCAAGCTCGGACTCTCTTGGACTTAC 540
DB 1126 AGCTCAGACGCTGATTCAGCCCATGGCTTTTCTTCCAAAGGAGATCATCACATTCGTC 1185
QY 541 AGG 543
DB 1186 AGG 1188
RESULT 15
US-09-529-063-26
; Sequence 26, Application US/09529063
; Patent No. US20020102542A1
; GENERAL INFORMATION:
; APPLICANT: FUKUSHIMA, DAIKICHI
; APPLICANT: SHIBAYAMA, SHIRO
; APPLICANT: TADA, HIDEAKI
; TITLE OF INVENTION: POLYPEPTIDE, cDNA ENCODING THE POLYPEPTIDE, AND USE OF
; FILE REFERENCE: Q58769
; CURRENT APPLICATION NUMBER: US/09/529,063
; CURRENT FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: PCT/JP98/04514
; PRIOR FILING DATE: 1998-10-06
; PRIOR APPLICATION NUMBER: JP 9-274674
; PRIOR FILING DATE: 1997-10-07
; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-529-063-26

Search completed: September 22, 2005, 22:20:25
Job time : 1517.01 secs

This Page Blank (uspto)